CANADIAN JOURNAL OF MATHEMATICS

Journal Canadien de Mathématiques

VOL. XI · NO. 3
1959

Published for
THE CANADIAN MATHEMATICAL CONGRESS
by the
University of Toronto Press

The chief languages of the *Journal* are English and French.

Manuscripts for publication in the *Journal* should be sent to the *Editor-in-Chief*, G. F. D. Duff, University of Toronto. Authors are asked to write with a sense of perspective and as clearly as possible, especially in the introduction. Regarding typographical conventions, attention is drawn to the Author's Manual of which a copy will be furnished on request.

All other correspondence should be addressed to the *Managing Editor*, G. de B. Robinson, University of Toronto.

The *Journal* is published quarterly. Subscriptions should be sent to the *Managing Editor*. The price per volume of four numbers is \$8.00. This is reduced to \$4.00 for individual members of recognized Mathematical Societies.

The Canadian Mathematical Congress gratefully acknowledges the assistance of the following towards the cost of publishing this Journal:

University of Alberta	Assumption University
University of British Columbia	Carleton College
Dalhousie University	Ecole Polytechnique
Université Laval	Loyola College
University of Manitoba	McGill University
McMaster University	Université de Montréal
Queen's University	Royal Military College
St. Mary's University	University of Toronto

National Research Council of Canada
and the
American Mathematical Society

AUTHORIZED AS SECOND CLASS MAIL, POST OFFICE DEPARTMENT, OTTAWA

r,
r,

ne
re
e,
s,
be

ng

nt
rs
d

e
l:

SUR LES REPRESENTATIONS UNITAIRES DES GROUPES DE LIE NILPOTENTS. IV

JACQUES DIXMIER

Soit n un entier $\geqslant 1$. On notera M_n l'ensemble des matrices carrées d'ordre n à éléments réels, et G_n le groupe des $x = (\xi_{jk}) \in M_n$ tels que $\xi_{jk} = 0$ pour $1 \leqslant j < k \leqslant n$, $\xi_{jj} = 1$ pour $1 \leqslant j \leqslant n$. Le groupe G_n est un groupe de Lie nilpotent simplement connexe, dont l'algèbre de Lie s'identifie à l'ensemble $\mathfrak{g}_n$ des $x = (\xi_{jk}) \in M_n$ tels que $\xi_{jk} = 0$ pour $1 \leqslant j \leqslant k \leqslant n$. Nous allons déterminer:

(1°) le centre de l'algèbre enveloppante de $\mathfrak{g}_n$;

(2°) la série "principale" de représentations unitaires irréductibles de G_n; (la recherche de *toutes* les représentations unitaires irréductibles de G_n ne semble pas facile);

(3°) la formule de Plancherel pour G_n;

(4°) les caractères globaux (au sens de **(5)**) des représentations de la série principale.

Pour n impair, l'étude est un peu plus compliquée que pour n pair. On exposera les démonstrations en détail pour n pair. Pour n impair, on insistera seulement sur les différences de calcul.

On emploiera les notations suivantes. La matrice $(\xi_{jk}) \in M_n$ telle que $\xi_{rs} = 1$ et $\xi_{jk} = 0$ pour $j \neq r$ ou $k \neq s$ sera notée e_{rs}. L'ensemble des matrices $(\xi_{jk}) \in M_n$ telles que $\xi_{jk} = 0$ pour $j + k \neq n + 1$ sera noté E_n. Pour toute matrice $x = (\xi_{jk}) \in M_n$, on posera:

$$\Delta_1(x) = \begin{vmatrix} \xi_{11} & \xi_{12} \dots \xi_{1n} \\ \xi_{21} & \xi_{22} \dots \xi_{2n} \\ \dots & \\ \xi_{n1} & \xi_{n2} \dots \xi_{nn} \end{vmatrix}$$

$$\Delta_2(x) = \begin{vmatrix} \xi_{12} & \xi_{13} & \dots \xi_{1n} \\ \xi_{22} & \xi_{23} & \dots \xi_{2n} \\ \dots & & \\ \xi_{n-1,2} & \xi_{n-1,3} & \dots \xi_{n-1,n} \end{vmatrix}$$

$$\dots\dots\dots\dots\dots\dots\dots\dots$$

$$\Delta_{n-1}(x) = \begin{vmatrix} \xi_{1,n-1} & \xi_{1n} \\ \xi_{2,n-1} & \xi_{2n} \end{vmatrix}$$

$$\Delta_n(x) = \xi_{1n}$$

$$\Delta_{n+1}(x) = 1.$$

Reçu le 15 août, 1958.

L'algèbre enveloppante d'une algèbre de Lie $\mathfrak{g}$ sera notée $\mathfrak{U}(\mathfrak{g})$. L'algèbre symétrique d'un espace vectoriel V sera notée $\mathfrak{S}(V)$. Sur le groupe G_n, la mesure définie par la forme différentielle $\Pi_{1\leqslant k<j\leqslant n} d\xi_{jk}$, qui est une mesure de Haar, sera appelée mesure de Haar canonique, et sera la seule utilisée. De même, sur le groupe additif des matrices (η_{jk}) à n lignes et n' colonnes, la seule mesure utilisée sera la mesure $\Pi_{1\leqslant j\leqslant n, 1\leqslant k\leqslant n'} d\eta_{jk}$.

Les lemmes 1 et 3 se trouvent dans (**3**, pp. 8-10 et 12-14). Toutefois, comme la situation est ici légèrement différente à certains égards, on a explicité les calculs pour la commodité du lecteur.

1. Cas où n est pair. Centre de $\mathfrak{U}(\mathfrak{g}_n)$. Nous poserons $n = 2m$. Tout $x \in G_n$ se met sous la forme

$$x = \begin{pmatrix} y & 0 \\ w & z \end{pmatrix}$$

où $y \in G_m$, $z \in G_m$, $w \in M_m$. Si

$$x' = \begin{pmatrix} y' & 0 \\ w' & z' \end{pmatrix}$$

on a

$$xx' = \begin{pmatrix} yy' & 0 \\ wy' + zw' & zz' \end{pmatrix}$$

d'où facilement

$$x^{-1} = \begin{pmatrix} y^{-1} & 0 \\ -z^{-1} w y^{-1} & z^{-1} \end{pmatrix} \qquad x x' x^{-1} = \begin{pmatrix} y y' y^{-1} & 0 \\ t & z z' z^{-1} \end{pmatrix}$$

avec $t = (wy' + zw' - zz'z^{-1}w)y^{-1}$. On voit que l'ensemble A_{2m} des $x \in G_{2m}$ tels que $y = z = 1$ est un sous-groupe distingué abélien de G_{2m}. L'idéal abélien $\mathfrak{a}_{2m}$ de $\mathfrak{g}_{2m}$ correspondant à A_{2m} est l'ensemble des matrices

$$\begin{pmatrix} 0 & 0 \\ w & 0 \end{pmatrix},$$

où $w \in M_m$.

LEMME 1. *Soit N_m l'ensemble des $w \in M_m$ tels que $\Delta_1(w)\Delta_2(w) \ldots \Delta_m(w) \neq 0$. Si $w \in N_m$, il existe des éléments $y \in G_m$, $z \in G_m$, $e \in E_m$ uniques tels que $w = zey$. Si $e = (\epsilon_{jk})$, on a*

$$\epsilon_{m-j+1,j} = (-1)^{m-j} \frac{\Delta_j(w)}{\Delta_{j+1}(w)}. \tag{1}$$

Démonstration. Posant $z^{-1} = z'$, il revient au même de prouver qu'il existe des éléments $y \in G_m$, $z' \in G_m$, $e \in E_m$ uniques tels que $z'w = ey$. Soient $w = (\omega_{jk})$, $y = (\eta_{jk})$, $z' = (\zeta_{jk})$, $e = (\epsilon_{jk})$. On doit avoir:

$$(2) \qquad \sum_{r=1}^{m} \zeta_{jr}\,\omega_{rk} = 0 \qquad (j+k > m+1)$$

$$(3) \qquad \sum_{r=1}^{m} \zeta_{jr}\,\omega_{r,m-j+1} = \epsilon_{j,m-j+1}$$

$$(4) \qquad \sum_{r=1}^{m} \zeta_{jr}\,\omega_{rk} = \epsilon_{j,m-j+1}\,\eta_{m-j+1,k} \qquad (j+k < m+1).$$

Pour j fixé, les équations (2), qui s'écrivent

$$\sum_{r=1}^{j-1} \zeta_{jr}\,\omega_{rk} = -\,\omega_{jk} \qquad (k = m-j+2, m-j+3, \ldots, m)$$

forment un système de $j-1$ équations à $j-1$ inconnues, dont le déterminant est $\Delta_{m-j+2}(w)$. Ce déterminant est non nul puisque $\omega \in N_m$. D'où l'existence et l'unicité des ζ_{jk} satisfaisant à (2). Les équations (3) donnent alors les $\epsilon_{j,m-j+1}$. D'ailleurs, en éliminant $\zeta_{j1}, \ldots, \zeta_{j,j-1}$ entre les j équations (2)-(3) qui contiennent ces inconnues,il vient

$$\begin{vmatrix} \omega_{1,m-j+1} & \omega_{1,m-j+2} \cdots \omega_{1,m} \\ \omega_{2,m-j+1} & \omega_{2,m-j+2} \cdots \omega_{2,m} \\ \cdots & \cdots \quad \cdots \\ \omega_{j,m-j+1} - \epsilon_{j,m-j+1} & \omega_{j,m-j+2} \cdots \omega_{j,m} \end{vmatrix} = 0.$$

D'où les formules (1).

Comme $w \in N_m$, on voit que $\epsilon_{1m} \neq 0$, $\epsilon_{2,m-1} \neq 0, \ldots, \epsilon_{m-1,2} \neq 0$. Les formules (4) prouvent alors l'existence et l'unicité des η_{jk}.

LEMME 2. *Soit $w = (\omega_{jk}) \to f(w)$ une fonction polynôme sur M_m. Pour qu'on ait $f(z\,w\,y) = f(w)$ quels que soient $w \in M_m$, $y \in G_m$, $z \in G_m$, il faut et il suffit que f soit dans l'algèbre engendrée par les fonctions* $\Delta_1, \Delta_2, \ldots, \Delta_m$.

Démonstration. Posant $w = (\omega_{jk})$, $y = (\eta_{jk})$, $z = (\zeta_{jk})$, $z\,w\,y = (\omega_{jk}')$, on a

$$\omega'_{jk} = \sum_{1\leq r\leq m,\, 1\leq s\leq m} \zeta_{jr}\,\omega_{rs}\,\eta_{sk} = \sum_{1\leq r\leq j,\, k\leq s\leq m} \zeta_{jr}\,\omega_{rs}\,\eta_{sk}.$$

Donc

$$(\omega'_{jk})_{1\leq j\leq t,\, m-t+1\leq k\leq m}$$
$$= (\zeta_{jk})_{1\leq j\leq t,\, 1\leq k\leq t}\,(\omega_{jk})_{1\leq j\leq t,\, m-t+1\leq k\leq m}\,(\eta_{jk})_{m-t+1\leq j\leq m,\, m-t+1\leq k\leq m}$$

et par suite

$$\det(\omega'_{jk})_{1\leq j\leq t,\, m-t+1\leq k\leq m} = \det(\omega_{jk})_{1\leq j\leq t,\, m-t+1\leq k\leq m}.$$

Ceci prouve que la condition de l'énoncé est suffisante.

Maintenant, soit $w = (\omega_{jk}) \to f(w) = f((\omega_{jk}))$ un polynôme tel que $f(z\,w\,y) = f(w)$ quels que soient $w \in M_m$, $y \in G_m$, $z \in G_m$. Si on remplace les ω_{jk} tels que $j + k \neq m + 1$ par 0 dans $f((\omega_{jk}))$, on obtient un polynôme par rapport

à $\omega_{m1}, \omega_{m-1,2}, \ldots, \omega_{1m}$, que nous noterons $g(\omega_{m1}, \ldots, \omega_{1m})$. Conservons la notation N_m du Lemme 1. Si $w \in N_m$, il existe $y \in G_m$, $z \in G_m$, $e = (\epsilon_{jk}) \in E_m$ tels que $w = z\,e\,y$. On a $f(w) = f(e) = g(\epsilon_{m1}, \ldots, \epsilon_{1m})$. D'après les formules (1),

$$f((\omega_{jk})) = g\left((-1)^{m+1}\frac{\Delta_1(w)}{\Delta_2(w)}, \ldots, -\frac{\Delta_{m-1}(w)}{\Delta_m(w)}, \Delta_m(w)\right). \tag{5}$$

Considérons maintenant les matrices w de la forme

$$\begin{pmatrix} \omega_{11} & \omega_{12} & \ldots & \omega_{1,m-1} & \omega_{1m} \\ 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & \ldots & 1 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \end{pmatrix}.$$

Quand on restreint f à l'ensemble de ces matrices, on obtient un polynôme $h(\omega_{11}, \ldots, \omega_{1m})$ et la formule (5) devient

$$h(\omega_{11}, \ldots, \omega_{1m}) = g\left(\pm\frac{\omega_{11}}{\omega_{12}}, \ldots, \frac{\omega_{1,m-1}}{\omega_{1,m}}, \omega_{1m}\right) \tag{6}$$

valable pour $\omega_{12} \neq 0, \omega_{13} \neq 0, \ldots, \omega_{1m} \neq 0$. Les égalités (5) et (6) entraînent

$$f((\omega_{jk})) = h(\pm\Delta_1(w), \ldots, -\Delta_{m-1}(w), \Delta_m(w)),$$

égalité valable pour $w \in N_m$ et par suite pour toute $w \in M_m$ d'après le principe d'inconséquence des inégalités algébriques. D'où le lemme.

THÉORÈME 1. *Le centre de* $\mathfrak{U}(\mathfrak{g}_{2m})$ *est engendré par les éléments algébriquement indépendants*

$$e_{2m,1}, \begin{vmatrix} e_{2m-1,1} & e_{2m-1,2} \\ e_{2m,1} & e_{2m,2} \end{vmatrix}, \ldots, \begin{vmatrix} e_{m+1,1} & \ldots & e_{m+1,m} \\ \ldots & \ldots & \ldots \\ e_{2m,1} & \ldots & e_{2m,m} \end{vmatrix}^{*}.$$

Démonstration. Nous allons d'abord chercher les éléments de $\mathfrak{S}(\mathfrak{g}_{2m})$ invariants pour la représentation adjointe de $\mathfrak{g}_{2m}$. Un élément de $\mathfrak{S}(\mathfrak{g}_{2m})$ est de la forme $f((e_{jk})_{j>k})$, où f est un polynôme en $\frac{1}{2}n(n-1)$ variables à coefficients réels. Les seuls crochets non nuls des e_{jk} entre eux sont données par les formules

$$[e_{jk}, e_{kl}] = e_{jl} = -[e_{kl}, e_{jk}] \qquad (j > k > l).$$

La condition que $f((e_{jk}))$ soit invariant pour la représentation adjointe se traduit par les égalités

*Les e_{jk} qui figurent dans ces déterminants appartiennent à $\mathfrak{a}_{2m}$, donc sont deux à deux permutables; ainsi, il n'y a pas d'ambiguïté sur la signification de ces déterminants.

$$\sum_{j>k} [e_{rs}, e_{jk}] f'_{e_{jk}} = 0 \qquad (r > s)$$

c'est-à-dire

$$(7) \qquad \sum_{k<s} e_{rk} f'_{e_{sk}} - \sum_{j>r} e_{js} f'_{e_{jr}} = 0 \qquad (r > s).$$

Cette égalité se réduit à $0 = 0$ pour $r = n$ et $s = 1$. Pour $r = n - 1$, $s = 1$, elle donne

$$e_{n1} f'_{e_{n,n-1}} = 0,$$

de sorte que f est indépendant de $e_{n,n-1}$. Pour $r = n$, $s = 2$, elle donne

$$e_{n1} f'_{e_{21}} = 0,$$

de sorte que f est indépendant de e_{21}. Soit p un entier $< m$, et supposons démontré que f est indépendant des e_{jk} pour $j \leqslant p$ d'une part, pour $k \geqslant n - p + 1$ d'autre part. Ecrivons la condition (7) pour $r = n - p$, $s = 1, 2, \ldots, p$ (ce qui est possible car $p < n - p$). Nous obtenons

$$\sum_{j>n-p} e_{j1} f'_{e_{j,n-p}} = 0$$

$$\sum_{j>n-p} e_{j2} f'_{e_{j,n-p}} = e_{n-p,1} f'_{e_{21}}$$

$$\ldots$$

$$\sum_{j>n-p} e_{jp} f'_{e_{j,n-p}} = e_{n-p,1} f'_{e_{p1}} + e_{n-p,2} f'_{e_{p2}} + \ldots + e_{n-p,p-1} f'_{e_{p,p-1}}.$$

D'après l'hypothèse de récurrence, les deuxièmes membres sont nuls. Ces égalités entraînent alors que

$$f'_{e_{j,n-p}} = 0$$

pour $j > n - p$, c'est-à-dire que f est indépendant des e_{jk} pour $k = n - p$. Ecrivant maintenant la condition (7) pour $s = p + 1$, $r = n - p + 1$, $n - p + 2, \ldots, n$ (ce qui est possible car $n - p + 1 > p + 1$), on trouve de même que f est indépendant des e_{jk} pour $j = p + 1$.

Ainsi, f est indépendant des e_{jk} pour $j \leqslant m$ d'une part, pour $k \geqslant m + 1$ d'autre part, de sorte que $f \in \mathfrak{S}(\mathfrak{a}_{2m})$. Cherchons donc les éléments de $\mathfrak{S}(\mathfrak{a}_{2m})$ invariants pour la représentation adjointe de $\mathfrak{g}_{2m}$, ou, ce qui revient au même, pour la représentation adjointe ρ de G_{2m}. Identifions $\mathfrak{a}_{2m}$ à M_m par l'application

$$\begin{pmatrix} 0 & 0 \\ w & 0 \end{pmatrix} \to w.$$

Alors, si

$$x = \begin{pmatrix} y & 0 \\ w' & z \end{pmatrix} \in G_{2m},$$

on a

$$\rho(x) \,.\, w = \rho(x) \,.\, \begin{pmatrix} 0 & 0 \\ w & 0 \end{pmatrix} = \begin{pmatrix} y & 0 \\ w' & z \end{pmatrix} \begin{pmatrix} 0 & 0 \\ w & 0 \end{pmatrix} \begin{pmatrix} y & 0 \\ w' & z \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 0 \\ z\, w\, y^{-1} & 0 \end{pmatrix} = z\, w\, y^{-1}.$$

Notons $A_{y,z}$ l'automorphisme de l'algèbre $\mathfrak{S}(M_m)$ qui prolonge l'automorphisme $w \to z\,w\,y^{-1}$ de l'espace vectoriel M_m. Il s'agit donc de trouver les éléments de $\mathfrak{S}(M_m)$ qui sont invariants pour les automorphismes $A_{y,z}$.

Grâce a la forme bilinéaire $(w, w') \to \mathrm{tr}(w\,w')$ sur M_m nous identifierons l'espace vectoriel M_m à son dual. Dans cette identification, e_{jk} s'identifie à la forme linéaire $(\omega_{jk}) \to \omega_{kj}$ sur M_m. Alors, $\mathfrak{S}(M_m)$ s'identifie à l'algèbre des fonctions polynômes sur M_m: à l'élément

$$e_{j_1k_1} \ldots e_{j_pk_p}$$

de $\mathfrak{S}(M_m)$ correspond la fonction polynôme

$$(\omega_{jk}) \to \omega_{k_1j_1} \ldots \omega_{k_pj_p}.$$

Pour $y \in G_m$, $z \in G_m$, on a

$$\begin{aligned} \mathrm{tr}(w(A_{y,z}w')) &= \mathrm{tr}(w\,z\,w'\,y^{-1}) = \mathrm{tr}(y^{-1}\,w\,z\,w') \\ &= \mathrm{tr}((A_{z^{-1},y^{-1}}\,w)w'). \end{aligned}$$

Donc le transposé de $A_{y,z}$ s'identifie à $A_{z^{-1},y^{-1}}$. Alors, d'après les propriétés élémentaires des algèbres symétriques, pour qu'un élément de $\mathfrak{S}(M_m)$ soit invariant par les $A_{y,z}$, il faut et il suffit que la fonction polynôme correspondante soit invariante par les $A_{y,z}$, c'est-à-dire par l'application $w \to z\,w\,y^{-1}$ de M_m sur M_m. Donc (Lemme 2) les éléments de $\mathfrak{S}(M_m)$ invariants pour les $A_{y,z}$ constituent l'algèbre engendrée par les éléments

$$e_{m,1},\quad \begin{vmatrix} e_{m-1,1} & e_{m-1,2} \\ e_{m,1} & e_{m,2} \end{vmatrix}, \ldots, \begin{vmatrix} e_{11} & \ldots & e_{1m} \\ \ldots & \ldots & \ldots \\ e_{m1} & \ldots & e_{mm} \end{vmatrix}.$$

Compte tenu de l'identification adoptée de $\mathfrak{a}_{2m}$ à M_m, on voit que les éléments de $\mathfrak{S}(\mathfrak{g}_{2m})$ invariants pour la représentation adjointe constituent l'algèbre engendrée par les éléments

$$e_{2m,1},\quad \begin{vmatrix} e_{2m-1,1} & e_{2m-1,2} \\ e_{2m,1} & e_{2m,2} \end{vmatrix}, \ldots, \begin{vmatrix} e_{m+1,1} & \ldots & e_{m+1,m} \\ \ldots & \ldots & \ldots \\ e_{2m,1} & \ldots & e_{2m,m} \end{vmatrix}.$$

Enfin, le centre de $\mathfrak{U}(\mathfrak{g}_{2m})$ est l'image, par l'application canonique ϕ de $\mathfrak{S}(\mathfrak{g}_{2m})$ sur $\mathfrak{U}(\mathfrak{g}_{2m})$, de l'algèbre précédente. Comme $\mathfrak{a}_{2m}$ est abélien, la restriction de ϕ à $\mathfrak{S}(\mathfrak{a}_{2m})$ est un isomorphisme de $\mathfrak{S}(\mathfrak{a}_{2m})$ sur $\mathfrak{U}(\mathfrak{a}_{2m}) \subset \mathfrak{U}(\mathfrak{g}_{2m})$. D'où le théorème.

2. Cas où n est pair. Formule de Plancherel. Nous conservons les notations précédentes. Tout élément $e \in E_m$ définit la forme linéaire $w \to \mathrm{tr}(e\,w)$ sur M_m, donc la forme linéaire

$$\begin{pmatrix} 0 & 0 \\ w & 0 \end{pmatrix} \to \mathrm{tr}(e\,w)$$

sur $\mathfrak{a}_{2m}$. D'autre part, l'application exponentielle de $\mathfrak{a}_{2m}$ sur A_{2m}, c'est-à-dire l'application

$$\begin{pmatrix} 0 & 0 \\ w & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ w & 1 \end{pmatrix}$$

est un isomorphisme du groupe additif $\mathfrak{a}_{2m}$ sur le groupe abélien A_{2m}, de sorte que l'application

$$\begin{pmatrix} 1 & 0 \\ w & 1 \end{pmatrix} \to \exp i \operatorname{tr}(e\, w) = \exp i \sum_{j=1}^{m} \epsilon_{j,m-j+1}\, \omega_{m-j+1,j}$$

(où $e = (\epsilon_{jk})$ et $w = (\omega_{jk})$) est un caractère ξ_e de A_{2m}. Nous noterons U_e la représentation unitaire de G_{2m} induite par ξ_e.

Tout élément de G_{2m} se met de manière unique sous la forme

$$\begin{pmatrix} 1 & 0 \\ w & 1 \end{pmatrix}\begin{pmatrix} y & 0 \\ 0 & z \end{pmatrix} = \begin{pmatrix} y & 0 \\ wy & z \end{pmatrix}$$

avec $y \in G_m$, $z \in G_m$, $w \in M_m$. Ainsi, G_{2m} est produit semi-direct de A_{2m} et d'un groupe canoniquement isomorphe à $G_m \times G_m$. En outre

$$(8) \qquad \begin{pmatrix} y' & 0 \\ 0 & z' \end{pmatrix}\begin{pmatrix} y & 0 \\ w & z \end{pmatrix} = \begin{pmatrix} y'y & 0 \\ z'w & z'z \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ z'wy^{-1}y'^{-1} & 1 \end{pmatrix}\begin{pmatrix} y'y & 0 \\ 0 & z'z \end{pmatrix}.$$

Par suite, l'espace hilbertien où opère U_e s'identifie canoniquement à $L_C^2(G_m \times G_m)$, $G_m \times G_m$ étant muni de sa mesure de Haar canonique et C désignant le corps complexe; et, si $(y', z') \to f(y', z')$ est un élément de $L_C^2(G_m \times G_m)$, la formule (8) prouve que, pour

$$x = \begin{pmatrix} y & 0 \\ w & z \end{pmatrix} \in G_{2m}$$

on a

$$(U_e(x)f)(y', z') = f(y'\,y, z'\,z)\, \xi_e\left(\begin{pmatrix} 1 & 0 \\ z'\,w\,y^{-1}\,y'^{-1} & 1 \end{pmatrix}\right)$$

ou encore

$$(9) \qquad (U_e(x)f)(y', z') = f(y'\,y, z'\,z) \exp i \operatorname{tr}(e\, z'\, w\, y^{-1}\, y'^{-1}).$$

La formule (9) définit explicitement la représentation U_e.

THÉORÈME 2. *Pour* $e = (\epsilon_{jk}) \in E_m$, *posons* $\epsilon_1 = \epsilon_{m,1}$, $\epsilon_2 = \epsilon_{m-1,2}, \ldots, \epsilon_m = \epsilon_{1,m}$.

(i) *La représentation* U_e *admet le caractère infinitésimal* χ_e (*au sens de* **(2)**) *défini par*

$$\chi_e(e_{2m,1}) = i\, \epsilon_m, \quad \chi_e\left(\begin{vmatrix} e_{2m-1,1} & e_{2m-1,2} \\ e_{2m,1} & e_{2m,2} \end{vmatrix}\right) = i^2\, \epsilon_{m-1}\, \epsilon_m, \ldots,$$

$$\chi_e\left(\begin{vmatrix} e_{m+1,1} & \ldots & e_{m+1,m} \\ \ldots & \ldots & \ldots \\ e_{2m,1} & \ldots & e_{2m,m} \end{vmatrix}\right) = i^m\, \epsilon_1\, \epsilon_2 \ldots \epsilon_m.$$

(ii) *Si on se limite aux $e \in E_m$ tels que $\epsilon_2, \epsilon_3, \ldots, \epsilon_m \neq 0$, les représentations U_e sont irréductibles et deux à deux inéquivalentes.*

(iii) *Si F est une fonction intégrable sur G_{2m}, on a*

$$\int_{G_{2m}} |F(x)|^2 dx = (2\pi)^{-m^2} \int \ldots \int \operatorname{tr}(U_e(F)^* U_e(F)) \epsilon_2^2 \epsilon_3^4 \ldots \epsilon_m^{2(m-1)} d\epsilon_1 d\epsilon_2 \ldots d\epsilon_m.$$

Démonstration. Si

$$x = \begin{pmatrix} 1 & 0 \\ w & 1 \end{pmatrix} \in A_{2m},$$

la formule (9) devient

$$(U_e(x)f)(y', z') = f(y', z') \exp i \operatorname{tr}(e\, z'\, w\, y'^{-1}).$$

Donc l'opérateur différentiel correspondant à l'élément

$$\begin{pmatrix} 0 & 0 \\ w & 0 \end{pmatrix}$$

de $\mathfrak{a}_{2m}$ est l'opérateur de multiplication par la fonction

$$(y', z') \to i \operatorname{tr}(e\, z'\, w\, y'^{-1}) = i \operatorname{tr}((y'^{-1}\, e\, z')w).$$

Pour y' et z' fixés, la valeur de cette fonction divisée par i est la forme linéaire sur $\mathfrak{a}_{2m}$ définie par $y'^{-1}\, e\, z'$. Cette forme linéaire se prolonge en un homomorphisme ψ de $\mathfrak{S}(\mathfrak{a}_{2m}) = \mathfrak{U}(\mathfrak{a}_{2m})$ dans le corps complexe. La valeur de ψ pour un élément invariant de $\mathfrak{S}(\mathfrak{a}_{2m})$ est indépendante de y' et z'. Ainsi, l'opérateur différentiel correspondant à un élément du centre de $\mathfrak{U}(\mathfrak{g}_{2m})$ est un opérateur scalaire. Donc la représentation U_e admet un caractère infinitésimal χ_e, et on a

$$\chi_e\left(\begin{vmatrix} e_{2m-j+1,1} & \ldots & e_{2m-j+1,j} \\ \ldots & \ldots & \ldots \\ e_{2m,1} & \ldots & e_{2m,j} \end{vmatrix} \right) = \begin{vmatrix} i \operatorname{tr}(e\, e_{m-j+1,1}) & \ldots & i \operatorname{tr}(e\, e_{m-j+1,j}) \\ \ldots & \ldots & \ldots \\ i \operatorname{tr}(e\, e_{m,1}) & \ldots & i \operatorname{tr}(e\, e_{m,j}) \end{vmatrix}$$

$$= \begin{vmatrix} 0 & 0 & \ldots & 0 & i\epsilon_{j,m-j+1} \\ 0 & 0 & \ldots & i\epsilon_{j-1,m-j+2} & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ i\epsilon_{1,m} & 0 & \ldots & 0 & 0 \end{vmatrix} = i^j \epsilon_{m-j+1} \epsilon_{m-j+2} \ldots \epsilon_m.$$

Ceci prouve (i).

Soit

$$x = \begin{pmatrix} y & 0 \\ w & z \end{pmatrix} \in G_{2m};$$

alors l'automorphisme intérieur de G_{2m} correspondant à x définit un automorphisme de A_{2m} donc de son dual. Nous allons montrer que, si $\epsilon_2, \epsilon_3, \ldots, \epsilon_m \neq 0$, x ne peut laisser fixe ξ_e que si $x \in A_{2m}$. Il en résultera (**(7)**; cf. aussi **(1)**) que U_e est irréductible. Or, dire que x laisse fixe ξ_e signifie que x laisse

fixe la forme linéaire $w' \to \operatorname{tr}(e\,w')$ sur M_m. Comme $\rho(x)\,.\,w' = z\,w'\,y^{-1}$, cela signifie encore que $\operatorname{tr}(e\,w') = \operatorname{tr}(e\,z\,w'\,y^{-1}) = \operatorname{tr}(y^{-1}\,e\,z\,w')$ quel que soit $w' \in M_m$, donc que $e = y^{-1}\,e\,z$, donc que $y = z = 1$ (Lemme 1). Ceci établit notre assertion.

Soit $e' = (\epsilon'_{jk}) \in E_m$, avec $\epsilon_2'\,\epsilon_3' \ldots \epsilon_m' \neq 0$ (où $\epsilon'_j = \epsilon'_{m-j+1,j}$). Si $e \neq e'$, (i) prouve que $\chi_e \neq \chi_{e'}$, donc que les représentations U_e, $U_{e'}$ sont inéquivalentes. Ainsi, (ii) est démontré.

Soit

$$x = \begin{pmatrix} y & 0 \\ w & z \end{pmatrix} \to F(x) = F_1(y, z, w)$$

une fonction intégrable sur G_{2m}. Pour $f, f' \in L_C^2\,(G_m \times G_m)$, on a

$$\begin{aligned}(U_e(F)f|f') &= \int_{G_{2m}} (U_e(x)f|f')\,F(x)\,dx \\ &= \int_{G_{2m}} F(x)\,dx \iint_{G_m \times G_m} f(y'y, z'\,z)\,\overline{f'(y', z')}\,\exp i\,\operatorname{tr}(e\,z'w\,y^{-1}y'^{-1})\,dy'\,dz'.\end{aligned}$$

La fonction

$$(x, y', z') \to F(x)\,f(y'\,y, z'\,z)\,\overline{f'(y', z')}\,\exp i\,\operatorname{tr}(e\,z'\,w\,y^{-1}\,y'^{-1})$$

sur $G_{2m} \times G_m \times G_m$ est mesurable pour $dx\,dy'\,dz'$, et

$$\begin{aligned}&\iiint^{*} |F(x)\,f(y'\,y, z'\,z)\,\overline{f(y', z')}\,\exp i\,\operatorname{tr}(e\,z'\,w\,y^{-1}\,y'^{-1})|\,dx\,dy'\,dz' \\ &\qquad = \int^{*} |F(x)|\,dx \iint^{*} |f(y'\,y, z'\,z)|\,|f'(y', z')|\,dy'\,dz' < +\infty.\end{aligned}$$

On peut donc appliquer le théorème de Lebesgue-Fubini qui donne

$$\begin{aligned}(10)\quad (U_e(F)f|f') &= \iiiint\!\!\int F_1(y, z, w) f(y'y, z'z)\overline{f'(y', z')}\,\exp i\,\operatorname{tr}(ez'wy^{-1}y'^{-1}) \\ &\qquad\qquad dydzdwdy'dz' \\ &= \iiiint\!\!\int F_1(y'^{-1}y, z'^{-1}z, w) f(y, z)\overline{f'(y', z')}\,\exp i\,\operatorname{tr}(ez'wy^{-1}) \\ &\qquad\qquad dydzdwdy'dz' \\ &= \iiiint f(y, z)\overline{f'(y', z')}\,[\int F_1(y'^{-1}y, z'^{-1}z, w)\,\exp i\,\operatorname{tr}(y^{-1}ez'\,w) \\ &\qquad\qquad dw]dydzdy'dz'.\end{aligned}$$

Donc ((**2**), Lemme 35)

$$\operatorname{tr}(U_e(F)^*U_e(F)) = \iiiint |\int F_1(y'^{-1}y, z'^{-1}z, w)\,\exp i\,\operatorname{tr}(y^{-1}ez'w)dw|^2\,dy\,dz\,dy'\,dz'$$

Comme on a identifié canoniquement l'espace vectoriel M_m à son dual, la transformée de Fourier de $(y, z, w) \to F_1(y, z, w)$ par rapport à la variable w est encore une fonction $(y, z, w) \to \bar{F}(y, z, w)$ sur $G_m \times G_m \times M_m$, et on a

$$\begin{aligned}(11)\quad \operatorname{tr}(U_e(F)^*\,U_e(F)) &= \iiiint |\bar{F}(y'^{-1}\,y, z'^{-1}\,z, y^{-1}\,e\,z')|^2\,dy\,dz\,dy'\,dz' \\ &= \iiiint |\bar{F}(y, z, y^{-1}\,y'^{-1}\,e\,z')|^2\,dy\,dz\,dy'\,dz' \\ &= \iiiint |\bar{F}(y, z, y'\,e\,z')|^2\,dy\,dz\,dy'\,dz'\end{aligned}$$

Pour achever la démonstration, nous aurons besoin d'un lemme. Adoptons la notation N_m du Lemme 1. Soit $F_m = E_m \cap N_m$, c'est-à-dire l'ensemble des $e = (\epsilon_{jk}) \in M_m$ tels que $\epsilon_{jk} = 0$ pour $j + k \neq m + 1$, $\epsilon_{1m}\, \epsilon_{2,m-1} \ldots \epsilon_{m-1,2} \neq 0$. Posons $\epsilon_{m,1} = \epsilon_1, \ldots, \epsilon_{1,m} = \epsilon_m$. Tout $w \in N_m$ s'écrit de manière unique sous la forme $w = z\, e\, y$, avec $y \in G_m$, $z \in G_m$, $e \in F_m$ (Lemme 1), d'où une bijection ϕ de N_m sur $G_m \times G_m \times F_m$. Alors :

LEMME 3. *La bijection ϕ transforme la mesure dw sur N_m en la mesure $dy\, dz\, de$, où*

$$de = \epsilon_2^2 \epsilon_3^4 \ldots \epsilon_m^{2(m-1)}\, d\epsilon_1\, d\epsilon_2 \ldots d\epsilon_m$$

Démonstration. Posons $w = (\omega_{jk})$, $y = (\eta_{jk})$, $e = (\epsilon_{jk})$, $z = (\zeta_{jk})$. Si $w = z\, e\, y$ on a

$$\begin{aligned}\omega_{jk} &= \sum_{r=1}^{m} \sum_{s=1}^{m} \zeta_{jr}\, \epsilon_{rs}\, \eta_{sk} = \sum_{r=1}^{m} \zeta_{jr}\, \epsilon_{r,m-r+1}\, \eta_{m-r+1,k} \\ &= \sum_{r \leqslant \inf(j,m-k+1)} \zeta_{jr}\, \epsilon_{m-r+1}\, \eta_{m-r+1,k}.\end{aligned}$$

D'où

$$\begin{aligned}d\omega_{jk} = \sum_{1 \leqslant r \leqslant \inf(j,m-k+1)} (\epsilon_{m-r+1}\, \eta_{m-r+1,k}\, d\zeta_{jr} &+ \zeta_{jr}\, \eta_{m-r+1,k}\, d\epsilon_{m-r+1} \\ &+ \zeta_{jr}\, \epsilon_{m-r+1}\, d\eta_{m-r+1,k}).\end{aligned}$$

On a en particulier $d\omega_{1m} = d\epsilon_m$. Supposons démontré que

$$\prod_{j \leqslant p,\, k \geqslant m-p+1} d\omega_{jk} = \pm\, \epsilon_{m-p+2}^2\, \epsilon_{m-p+3}^4 \ldots \epsilon_m^{2(p-1)} \Big(\prod_{1 \leqslant k < j \leqslant p} d\zeta_{jk}\Big)\Big(\prod_{m-p+1 \leqslant j \leqslant m} d\epsilon_j\Big) \Big(\prod_{m-p+1 \leqslant k < j \leqslant m} d\eta_{jk}\Big). \tag{12}$$

On a alors

$$\begin{aligned}&\Big(\prod_{j \leqslant p,\, k \geqslant m-p+1} d\omega_{jk}\Big)\, d\omega_{1,m-p} \\ &= \pm\, \epsilon_{m-p+2}^2\, \epsilon_{m-p+3}^4 \ldots \epsilon_m^{2(p-1)} \Big(\prod_{1 \leqslant k < j \leqslant p} d\zeta_{jk}\Big)\Big(\prod_{m-p+1 \leqslant j \leqslant m} d\epsilon_j\Big)\Big(\prod_{m-p+1 \leqslant k < j \leqslant m} d\eta_{jk}\Big) \\ &\qquad (\eta_{m,m-p}\, d\epsilon_m + \epsilon_m\, d\eta_{m,m-p}) \\ &= \pm\, \epsilon_{m-p+2}^2\, \epsilon_{m-p+3}^4 \ldots \epsilon_m^{2(p-1)} \Big(\prod_{1 \leqslant k < j \leqslant p} d\zeta_{jk}\Big)\Big(\prod_{m-p+1 \leqslant j \leqslant m} d\epsilon_j\Big)\Big(\prod_{m-p+1 \leqslant k < j \leqslant m} d\eta_{jk}\Big) \\ &\qquad \epsilon_m d\eta_{m\ m-p}.\end{aligned}$$

Supposons démontré, pour un entier q tel que $1 \leqslant q \leqslant p$, que

$$\begin{aligned}&\Big(\prod_{j \leqslant p,\, k \geqslant m-p+1} d\omega_{jk}\Big)\, d\omega_{1,m-p}\, d\omega_{2,m-p} \ldots d\omega_{q-1,m-p} \\ &= \pm\, \epsilon_{m-p+2}^2\, \epsilon_{m-p+3}^4 \ldots \epsilon_m^{2(p-1)} \Big(\prod_{1 \leqslant k < j \leqslant p} d\zeta_{jk}\Big)\Big(\prod_{m-p+1 \leqslant j \leqslant m} d\epsilon_j\Big)\Big(\prod_{m-p+1 \leqslant k < j \leqslant m} d\eta_{jk}\Big) \\ &\qquad \epsilon_m\, \epsilon_{m-1} \ldots \epsilon_{m-q+2}\, d\eta_{m\ m-p}\, d\eta_{m-1,m-p} \ldots d\eta_{m-q+2,m-p}.\end{aligned} \tag{13}$$

Alors

$$(14)\quad \Big(\prod_{j\leq p,\, k\geq m-p+1} d\omega_{jk}\Big)\, d\omega_{1,m-p}\, d\omega_{2,m-p} \ldots d\omega_{q,m-p}$$
$$= \pm\, \epsilon_{m-p+2}^{2}\, \epsilon_{m-p+3}^{4} \ldots \epsilon_{m}^{2(p-1)} \Big(\prod_{1\leq k<j\leq p} d\zeta_{jk}\Big)\Big(\prod_{m-p+1\leq j\leq m} d\epsilon_j\Big)\Big(\prod_{m-p+1\leq k<j\leq m} d\eta_{jk}\Big)$$
$$\epsilon_m \epsilon_{m-1} \ldots \epsilon_{m-q+2}\, d\eta_{m,m-p} \ldots d\eta_{m-q+2,m-p}(\epsilon_{m-q+1}\, d\eta_{m-q+1,m-p})$$

Le passage de (13) à (14) prouve, par récurrence sur q, que

$$\Big(\prod_{j\leq p,\, k\geq m-p+1} d\omega_{jk}\Big)\, d\omega_{1,m-p} \ldots d\omega_{p,m-p}$$
$$= \pm\, \epsilon_{m-p+2}^{2} \ldots \epsilon_{m}^{2(p-1)} \epsilon_{m-p+1} \ldots \epsilon_m \Big(\prod_{1\leq k<j\leq p} d\zeta_{jk}\Big)\Big(\prod_{m-p+1\leq j\leq m} d\epsilon_j\Big)\Big(\prod_{m-p\leq k<j\leq m} d\eta_{jk}\Big).$$

On passe de même de là, par récurrence, à la formule

$$\Big(\prod_{j\leq p,\, k\geq m-p+1} d\omega_{jk}\Big) d\omega_{1,m-p} \ldots d\omega_{p,m-p}\, d\omega_{p+1,m} \ldots d\omega_{p+1,m-p+1}$$
$$= \pm\, \epsilon_{m-p+1}^{2} \epsilon_{m-p+2}^{4} \ldots \epsilon_{m}^{2p} \Big(\prod_{1\leq k<j\leq p+1} d\zeta_{jk}\Big)\Big(\prod_{m-p+1\leq j\leq m} d\epsilon_j\Big)\Big(\prod_{m-p\leq k<j\leq m} d\eta_{jk}\Big).$$

Enfin, en multipliant par $d\omega_{p+1,m-p}$, il vient

$$(15)\quad \prod_{j\leq p+1,\, k\geq m-p} d\omega_{jk} = \pm\, \epsilon_{m-p+1}^{2} \ldots \epsilon_{m}^{2p} \Big(\prod_{1\leq k<j\leq p+1} d\zeta_{jk}\Big)\Big(\prod_{m-p\leq j\leq m} d\epsilon_j\Big)$$
$$\Big(\prod_{m-p\leq k<j\leq m} d\eta_{jk}\Big).$$

Le passage de (12) à (15) prouve, par récurrence sur p, le Lemme 3.

Ceci posé, revenons à la formule (11). Elle entraîne

$$\int_{F_m} \operatorname{tr}(U_e(F)^* U_e(F))\, de = \int_{G_m}\int_{G_m}\int_{G_m}\int_{G_m}\int_{F_m} |\bar{F}(y, z, y'ez')|^2\, dy\, dz\, dy'\, dz'\, de$$
$$= \int_{G_m}\int_{G_m}\int_{N_m} |\bar{F}(y, z, w)|^2\, dy\, dz\, dw = \int_{G_m}\int_{G_m}\int_{M_m} |\bar{F}(y, z, w)|^2\, dy\, dz\, dw.$$

D'où, utilisant la formule de Plancherel sur le groupe abélien M_m,

$$\int_{E_m} \operatorname{tr}(U_e(F)^* U_e(F))\, de = (2\pi)^{m^2} \int_{G_m}\int_{G_m}\int_{M_m} |F_1(y, x, w)|^2\, dy\, dz\, dw$$
$$= (2\pi)^{m^2} \int_{G_{2m}} |F(x)|^2\, dx.$$

Ceci achève la démonstration du Théorème 2.

3. Cas où n est pair. Caractères globaux des représentations U_e.

LEMME 4. *Soient $P_1(x_1, \ldots, x_r), \ldots, P_s(x_1, \ldots, x_r)$ des polynômes à coefficients réels. Soit φ l'application de R^r dans R^s définie par les égalités $y_1 = P_1(x_1, \ldots, x_r), \ldots, y_s = P_s(x_1, \ldots, x_r)$. On suppose qu'il existe des constantes $A > 0$, $\delta > 0$, telles que $\|\varphi(x)\| \geq A\|x\|^\delta$ pour $x \in R^r$, $\|x\| \geq 1$ (on pose $\|x\| = |x_1| + \ldots + |x_r|$, $\|y\| = |y_1| + \ldots + |y_s|$ pour $x = (x_1, \ldots, x_r)$*

$\in R^r$, $y = (y_1, \ldots, y_s) \in R^s$). *Alors, si* $f \in \mathscr{S}(R^s)$ *(avec les notations de* **(8)**), *on a* $f\circ\varphi \in \mathscr{S}(R^r)$, *et l'application* $f \to f\circ\varphi$ *de* $\mathscr{S}(R^s)$ *dans* $\mathscr{S}(R^r)$ *est continue.*

Démonstration. Soit f une fonction numérique quelconque sur R^s. On a, pour tout $\alpha > 0$,

$$\sup_{x\in R^r, ||x||>1}|(f\circ\phi)(x)|\ ||x||^\alpha < A^{-\alpha/\beta} \sup_{x\in R^r, ||x||>1}|f(\varphi(x))|\ ||\varphi(x)||^{\alpha/\beta}$$
$$< A^{-\alpha/\beta} \sup_{y\in R^s}|f(y)|\ ||y||^{\alpha/\beta}.$$

Supposons maintenant que $f \in \mathscr{S}(R^s)$. Alors, $f\circ\varphi$ est indéfiniment dérivable sur R^r. Toute dérivée partielle $D(f\circ\varphi)$ de $f\circ\varphi$ est somme de termes de la forme $Q\,.\,((D'f)\circ\varphi)$ où Q est un polynôme et où D' est une dérivation partielle (on le voit aussitôt par récurrence sur l'ordre de D). Il existe des constantes $B > 0$, $\epsilon > 0$ telles que $||Q(x)|| < B||x||^\epsilon$ pour $||x|| > 1$. Alors

$$\sup_{x\in R^r, ||x||>1}|Q(x)((D'f)\circ\phi)(x)|\ ||x||^\alpha < B \sup_{x\in R^r, ||x||>1}|((D'f)\circ\phi)(x)|\ ||x||^{\alpha+\epsilon}$$
$$< BA^{-\alpha/\beta} \sup_{y\in R^s}|(D'f)(y)|\ ||y||^{(\alpha+\epsilon)/\beta}$$

D'autre part,

$$\sup_{x\in R^r, ||x||<1}|Q(x)((D'f)\circ\varphi)(x)|\ ||x||^\alpha < (\sup_{x\in R^r, ||x||<1}|Q(x)|)$$
$$\sup_{y\in R^s}|(D'f)(y)|.$$

Ceci prouve le lemme, en utilisant la définition de la topologie de $\mathscr{S}(R^r)$ et $\mathscr{S}(R^s)$ par semi-normes.

Soit maintenant $x = (\xi_{jk}) \to F(x) = F((\xi_{jk}))$ une fonction sur G_n. (Rappelons que $\xi_{jk} = 0$ pour $j < k$ et que $\xi_{jj} = 1$ si $x \in G_n$). Dire que F est indéfiniment dérivable sur G_n revient à dire que F est une fonction indéfiniment dérivable des ξ_{jk} ($j > k$). Si de plus F est une fonction indéfiniment dérivable à décroissance rapide des ξ_{jk} ($j > k$), on dira que F est indéfiniment dérivable à décroissance rapide sur G_n. Il revient au même de dire que F, transportée sur $\mathfrak{g}_n$ grâce à l'application exponentielle (qui est un isomorphisme de la variété différentiable $\mathfrak{g}_n$ sur la variété différentiable G_n), devient une fonction indéfiniment dérivable á décroissance rapide au sens usuel sur l'espace vectoriel $\mathfrak{g}_n$.

THÉORÈME 3. *Soit* $e = (\epsilon_{jk}) \in E_m$; *posons* $\epsilon_1 = \epsilon_{m,1}, \ldots, \epsilon_m = \epsilon_{1,m}$; *supposons* $\epsilon_2\epsilon_3 \ldots \epsilon_m \neq 0$.

(i) *Soit*

$$x = \begin{pmatrix} y & 0 \\ w & z \end{pmatrix} \to F(x) = F_1(y, z, w)$$

une fonction indéfiniment dérivable à décroissance rapide sur G_{2m}. *Alors, l'opérateur* $U_e(F)$ *est un opérateur à trace.*

(ii) *Il existe une distribution tempérée* T_e *sur* M_m *telle que*

$$\operatorname{tr}(U_e(F)) = \int F_1(1, 1, w)dT_e(w).$$

(iii) *Sur l'ensemble* N_m *des* $w \in M_m$ *tels que* $\Delta_2(w)\Delta_3(w) \ldots \Delta_m(w) \neq 0$, T_e *coïncide avec la fonction*

$$w \to (2\pi)^{\frac{1}{2}m(m-1)}|\epsilon_2^{-1}\epsilon_3^{-2}\dots\epsilon_m^{-m+1}|$$
$$\frac{\exp i\left[\epsilon_1\Delta_m(w) - \epsilon_2\dfrac{\Delta_{m-1}(w)}{\Delta_m(w)} + \epsilon_3\dfrac{\Delta_{m-2}(w)}{\Delta_{m-1}(w)} - \dots + (-1)^{m+1}\epsilon_m\dfrac{\Delta_1(w)}{\Delta_2(w)}\right]}{|\Delta_2(w)\Delta_3(w)\dots\Delta_m(w)|}.$$

Démonstration. Comme dans la démonstration du Théorème 2, introduisons la fonction $(y, z, w) \to \bar{F}(y, z, w)$ sur $G_m \times G_m \times M_m$, transformée de Fourier de la fonction $(y, z, w) \to F_1(y, z, w)$ par rapport à la variable w. La formule (10) prouve que $U_e(F)$ est défini par un noyau $(y, z, y', z') \to K(y, z, y', z')$ sur $(G_m \times G_m) \times (G_m \times G_m)$, ce noyau étant donné par la formule

$$K(y, z, y', z') = \bar{F}(y'^{-1}y, z'^{-1}z, y^{-1}ez').$$

Nous allons montrer qu'il existe des constantes $A > 0$ et $\delta > 0$ telles que

$$(16) \qquad ||y'^{-1}y|| + ||z'^{-1}z|| + ||y^{-1}ez'|| > A(||y|| + ||z|| + ||y'|| + ||z'||)^\delta$$

(où l'on pose $||y|| = \Sigma_{j>k}|\eta_{jk}|$ pour un élément $y = (\eta_{jk})$ de G_m, et $||w|| = \Sigma|\omega_{jk}|$ pour un élément $w = (\omega_{jk})$ de M_m). Il est immédiat qu'il existe des constantes $A_1 > 0$, $A_2 > 0$, $\delta_1 > 0$, $\delta_2 > 0$ telles que

$$||y^{-1}|| < A_1||y||^{\delta_1},\ ||yz|| < A_2(||y|| + ||z||)^{\delta_2}$$

quels que soient $y \in G_m$, $z \in G_m$ avec $||y|| > 1$, $||z|| > 1$. D'autre part, la démonstration du Lemme 1 prouve qu'il existe des constantes $A_3 > 0$, $\delta_3 > 0$ telles que

$$||y|| < A_3||y^{-1}ez'||^{\delta_3}$$

quels que soient $y \in G_m$, $z' \in G_m$. (Rappelons que $e \in E_m$ est fixé et que $\epsilon_2 \dots \epsilon_m \neq 0$; on tiendra compte aussi du fait que $||y^{-1}ez'|| > |\epsilon_m| > 0$). Prenant les transposés, on en déduit l'existence de constantes $A_4 > 0$, $\delta_4 > 0$ telles que

$$||z'|| < A_4||y^{-1}ez'||^{\delta_4}$$

quels que soient $y \in G_m$, $z' \in G_m$. D'où l'existence de constantes $A_5 > 0$, $\delta_5 > 0$ telles que

$$||y'|| = ||((y'^{-1}y)y^{-1})^{-1}|| < A_5(||y'^{-1}y|| + ||y^{-1}ez'||)^{\delta_5},$$
$$||z|| = ||(z'(z'^{-1}z))|| < A_5(||z'^{-1}z|| + ||y^{-1}ez'||)^{\delta_5}$$

quels que soient $y, y', z, z' \in G_m$. On en déduit enfin l'existence de constantes $A_6 > 0$, $\delta_6 > 0$ avec

$$||y|| + ||z|| + ||y'|| + ||z'|| < A_6(||y'^{-1}y|| + ||z'^{-1}z|| + ||(y^{-1}ez'||)^{\delta_6}$$

quels que soient $y, y', z, z' \in G_m$. D'où (16).

Comme F est une fonction indéfiniment dérivable à décroissance rapide sur $G_m \times G_m \times M_m$, le Lemme 4 prouve alors que K est une fonction indéfiniment dérivable à décroissance rapide sur $(G_m \times G_m) \times (G_m \times G_m)$. L'opérateur

$U_e(F)$, défini par le noyau K, est donc un opérateur d'Hilbert-Schmidt. Posons $X = G_m \times G_m$, de sorte que $K \in \mathscr{S}(X \times X)$ avec les notations de **(8)**. L'application canonique de $\mathscr{S}(X)$ dans $L^2(X)$ (il s'agit de l'espace L^2 pour la mesure de Haar canonique de $G_m \times G_m$) définit une application canonique de $\mathscr{S}(X) \hat{\otimes} \mathscr{S}(X)$ dans $L^2(X) \hat{\otimes} L^2(X)$ (cf. **(6)** pour les notations utilisées ici concernant les produits tensoriels topologiques). Or, comme $\mathscr{S}(X)$ est un espace nucléaire, on a $\mathscr{S}(X) \hat{\otimes} \mathscr{S}(X) = \mathscr{S}(X \times X)$, d'où une application canonique $\lambda : \mathscr{S}(X \times X) \to L^2(X) \hat{\otimes} L^2(X)$; d'autre part $L^2(X) \hat{\otimes} L^2(X)$ s'identifie à l'ensemble des opérateurs à trace dans $L^2(X)$, et il est immédiat que $\lambda(K)$ n'est autre que l'opérateur défini par K, c'est-à-dire $U_e(F)$. D'où (i). Par ailleurs, si $K_1 \in \mathscr{S}(X \times X)$, on a

$$\mathrm{tr}(\lambda(K_1)) = \int K_1(\xi, \xi)\, d\xi;$$

c'est évident si K_1 est un noyau élémentaire de la forme $(\xi, \xi') \to \varphi(\xi)\varphi'(\xi')$, où $\varphi \in \mathscr{S}(X)$, $\varphi' \in \mathscr{S}(X)$; et le cas général se déduit de là par linéarité et continuité. Donc

$$\mathrm{tr}(U_e(F)) = \iint K(y, z, y, z)\, dy\, dz = \iint \bar{F}(1, 1, yez)\, dy\, dz. \tag{17}$$

Or l'application $(y, z) \to yez$ de $G_m \times G_m$ dans M_m transforme la mesure $dy\, dz$ en une distribution tempérée sur M_m (à cause du Lemme 4 et des inégalités

$$||y|| < A_3||y^{-1}ez||^{s_3},\ ||z|| < A_4||y^{-1}ez||^{s_4});$$

soit T_e la transformée de Fourier de cette distribution, qui est aussi une distribution tempérée; on a

$$\mathrm{tr}(U_e(F)) = \int F_1(1, 1, w)\, dT_e(w).$$

D'où (ii).

On a

$$\begin{aligned} \bar{F}(1, 1, yez) &= \int F_1(1, 1, w) \exp i\, \mathrm{tr}(wyez)\, dw \\ &= \int F_1(1, 1, wy^{-1}) \exp\, i\, \mathrm{tr}(wez)\, dw. \end{aligned}$$

Posons $z = 1 + z^\circ$, de sorte que z° parcourt l'espace vectoriel $\mathfrak{g}_m$. On a alors

$$\begin{aligned} \bar{F}(1, 1, yez) &= \int F_1(1, 1, wy^{-1}) \exp i\, \mathrm{tr}(we)\, .\, \exp i\, \mathrm{tr}(wez^\circ)\, dw \\ &= \int \Phi(w) \exp i\, \mathrm{tr}(wez^\circ)\, dw \end{aligned}$$

en posant $\Phi(w) = F_1(1, 1, wy^{-1}) \exp i\, \mathrm{tr}(we)$. La fonction $w \to \Phi(w)$ est indéfiniment dérivable à décroissance rapide sur M_m. Soient $w = (\omega_{jk})$, $z = (\zeta_{jk})$. On a

$$\begin{aligned} \mathrm{tr}(wez^0) &= \sum_{j+k<m+1} \omega_{jk}\, \epsilon_{k, m-k+1}\, \zeta_{m-k+1, j} \\ &= \sum_{j+k<m+1} \omega_{jk}\, \epsilon_{m-k+1}\, \zeta_{m-k+1, j}. \end{aligned}$$

Donc

$$\bar{F}(1, 1, yez) = \int \dots \int \Phi((\omega_{jk})_{1\leq j\leq m,\, 1\leq k\leq m})\Big(\exp i \sum_{j+k<m+1} \omega_{jk}\epsilon_{m-k+1}\zeta_{m-k+1,\, j}\Big) d\omega_{11}\dots d\omega_{mm}$$
$$= \int \dots \int \Psi((\omega_{jk})_{j+k<m+1})\Big(\exp i \sum_{j+k<m+1} \omega_{jk}\,\epsilon_{m-k+1}\,\zeta_{m-k+1,\, j}\Big) \prod_{j+k<m+1} d\omega_{jk}$$

avec

$$\Psi((\omega_{jk})_{j+k<m+1}) = \int \dots \int \Phi((\omega_{jk})_{1\leq j\leq m,\, 1\leq k\leq m}) \prod_{j+k\geq m+1} d\omega_{jk}.$$

La fonction Ψ est indéfiniment dérivable à décroissance rapide, et l'on a

$$\int \bar{F}(1, 1, yez)dz$$
$$= \int \dots \int \prod_{j+k<m+1} d\zeta_{m-k+1,\, j} \int \dots \int \Psi((\omega_{jk})_{j+k<m+1})\Big(\exp i \sum_{j+k<m+1} \omega_{jk}\,\epsilon_{m-k+1}\,\zeta_{m-k+1,\, j}\Big) \prod_{j+k<m+1} d\omega_{jk}$$
$$= |\epsilon_2^{-1}\epsilon_3^{-2}\dots\epsilon_m^{-m+1}| \int \dots \int \prod_{j+k<m+1} d\zeta_{m-k+1,\, j} \int \dots \int \Psi((\omega_{jk})_{j+k<m+1}) \Big(\exp i \sum_{j+k<m+1} \omega_{jk}\,\zeta_{m-k+1,\, j}\Big) \prod_{j+k<m+1} d\omega_{jk}$$
$$= (2\pi)^{m(m-1)/2}|\epsilon_2^{-1}\epsilon_3^{-2}\dots\epsilon_m^{-m+2}|\, \Psi(0)$$
$$= (2\pi)^{m(m-1)/2}|\epsilon_2^{-1}\epsilon_3^{-2}\dots\epsilon_m^{-m+1}| \int \Phi(t)dt$$

où l'on fait varier t dans l'ensemble $\mathfrak{h}_m$ des matrices $(\tau_{jk}) \in \dot{M}_m$ telles que $\tau_{jk} = 0$ pour $j + k < m + 1$, et où dt désigne la mesure définie par la forme différentielle

$$\prod_{j+k\geq m+1} d\tau_{jk}$$

Ainsi

$$\int \bar{F}(1, 1, ye\, z)dz = (2\pi)^{\frac{1}{2}m(m-1)}|\epsilon_2^{-1}\epsilon_3^{-2}\dots\epsilon_m^{-m+1}| \int F_1(1, 1, ty^{-1})\exp i\, \mathrm{tr}(te)\, dt.$$

Toute matrice $t = (\tau_{jk}) \in \mathfrak{h}_m$ telle que $\tau_{1,m}\tau_{2,m-1}\dots\tau_{m,1} \neq 0$ se met de manière unique sous la forme ze', où $e' = (\epsilon_{jk}') \in E_m$ et où $z = (\zeta_{jk}) \in G_m$. Posant $\epsilon'_{m-j+1,\, j} = \epsilon_j'$, on a $\tau_{jk} = \zeta_{j,m-k+1}\,\epsilon_k'$, d'où $\epsilon_k' = \tau_{m-k+1,k}$. Donc

$$\mathrm{tr}(te) = \sum_j \tau_{j,m-j+1}\,\epsilon_j = \sum_j \epsilon'_{m-j+1}\,\epsilon_j$$
$$dt = |\epsilon_2'\epsilon_3'^2\dots\epsilon_m'^{m-1}|\, d\epsilon_1'\dots d\epsilon_m'\, dz$$

ou, compte tenu du Lemme 3,

$$dt = |\epsilon_2'^{-1}\epsilon_3'^{-2}\dots\epsilon_m'^{-m+1}|\, de'dz.$$

Par suite

$$\int \bar{F}(1, 1, yez)dz = (2\pi)^{\frac{1}{2}m(m-1)}|\epsilon_2^{-1}\epsilon_3^{-2} \ldots \epsilon_m^{-m+1}| \iint F_1(1, 1, ze'y^{-1})$$
$$\left(\exp i \sum_j \epsilon'_{m-j+1}\, \epsilon_j\right) |\epsilon_2'^{-1}\epsilon_3'^{-2} \ldots \epsilon_m'^{-m+1}|\, de'dz.$$

Soit F_2 la fonction $w \to F_1\ (1, 1, w)$ sur M_m. Supposons maintenant que le support de F_2 soit compact et contenu dans N_m. Alors, la fonction $w \to F_2(w)|\Delta_2(w)^{-1}\Delta_3(w)^{-1} \ldots \Delta_m(w)^{-1}|$ est intégrable pour dw. Donc la fonction

$$(z, e', y) \to F_2(ze'y)|\epsilon_2'^{-1}\epsilon_3'^{-2} \ldots \epsilon_m'^{-m+1}|$$

est intégrable pour la mesure $dy\, de'\, dz$. Donc

$$\iint \bar{F}(1, 1, yez)\, dydz = (2\pi)^{\frac{1}{2}m(m-1)}|\epsilon_2^{-1}\epsilon_3^{-2} \ldots \epsilon_m^{-m+1}| \iiint F_1(1, 1, ze'y^{-1})$$
$$\left(\exp i \sum_j \epsilon'_{m-j+1}\, \epsilon_j\right) |\epsilon_2'^{-1}\epsilon_3'^{-2} \ldots \epsilon_m'^{-m+1}|\, dyde'dz$$
$$= (2\pi)^{\frac{1}{2}m(m-1)}|\epsilon_2^{-1}\epsilon_3^{-2} \ldots \epsilon_m^{-m+1}| \int F_1(1, 1, w)$$
$$\frac{\exp i\left[\epsilon_1\Delta_m(w) - \epsilon_2\dfrac{\Delta_{m-1}(w)}{\Delta_m(w)} + \ldots + (-1)^{m+1}\epsilon_m\dfrac{\Delta_1(w)}{\Delta_2(w)}\right]}{|\Delta_2(w)\Delta_3(w) \ldots \Delta_m(w)|}\, dw$$

ce qui prouve (iii).

Remarque. Sauf dans le cas trivial où $m = 1$, la fonction $w \to \Delta_2(w)^{-1} \ldots \Delta_m(w)^{-1}$ est non localement intégrable pour dw, de sorte que T_e n'est pas une mesure.

4. Cas où n est impair. Centre de $\mathfrak{U}(\mathfrak{g}_n)$. Nous poserons alors $n = 2m+1$. Tout élément x de G_{2m+1} se met sous la forme

$$x = \begin{pmatrix} y & 0 & 0 \\ u & 1 & 0 \\ w & v & z \end{pmatrix}$$

où $y \in G_m$, $z \in G_m$, $w \in M_m$, et où u (resp. v) est une matrice à 1 ligne et m colonnes (resp. 1 colonne et m lignes). Si

$$x' = \begin{pmatrix} y' & 0 & 0 \\ u' & 1 & 0 \\ w' & v' & z' \end{pmatrix}$$

on a

$$xx' = \begin{pmatrix} y\,y' & 0 & 0 \\ u\,y' + u' & 1 & 0 \\ w\,y' + v\,u' + z\,w' & v + z\,v' & z\,z' \end{pmatrix}.$$

D'où facilement

$$x^{-1} = \begin{pmatrix} y^{-1} & 0 & 0 \\ -u\,y^{-1} & 1 & 0 \\ z^{-1}(vu - w)y^{-1} & -z^{-1}v & z^{-1} \end{pmatrix}$$

et

$$x\,x'\,x^{-1} = \begin{pmatrix} y\,y'\,y^{-1} & 0 & 0 \\ (u\,y' + u' - u)y^{-1} & 1 & 0 \\ t & v + z\,v' - z\,z'\,z^{-1}v & z\,z'\,z^{-1} \end{pmatrix} \tag{18}$$

avec $t = (w\,y' + v\,u' + z\,w' - v\,u - z\,v'\,u + z\,z'\,z^{-1}\,v\,u - z\,z'\,z^{-1}\,w)y^{-1}$. On voit que l'ensemble A_{2m+1} des $x \in G_{2m+1}$ tels que $y = z = 1$, $u = v = 0$ est un sous-groupe distingué abélien de G_{2m+1}. L'idéal abélien $\mathfrak{a}_{2m+1}$ de $\mathfrak{g}_{2m+1}$ correspondant à A_{2m+1} est l'ensemble des matrices de la forme

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ w & 0 & 0 \end{pmatrix}.$$

THÉORÈME 4. *Le centre de $\mathfrak{U}(\mathfrak{g}_{2m+1})$ est engendré par les éléments algébriquement indépendants*

$$e_{2m+1,1}, \quad \begin{vmatrix} e_{2m,1} & e_{2m,2} \\ e_{2m+1,1} & e_{2m+1,2} \end{vmatrix}, \quad \ldots, \quad \begin{vmatrix} e_{m+2,1} & \ldots & e_{m+2,m} \\ \ldots & & \ldots\ldots \\ e_{2m+1,1} & \ldots & e_{2m+1,m} \end{vmatrix}.$$

Démonstration. Comme dans la démonstration du Théorème 1, on prouve d'abord qu'un élément de $\mathfrak{S}(\mathfrak{g}_{2m+1})$ invariant pour la représentation adjointe est dans $\mathfrak{S}(\mathfrak{a}_{2m+1})$. Soit ρ la représentation adjointe de G_{2m+1} dans $\mathfrak{a}_{2m+1}$. Identifions $\mathfrak{a}_{2m+1}$ à M_m par l'application

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ w & 0 & 0 \end{pmatrix} \to w.$$

Alors, si

$$x = \begin{pmatrix} y & 0 & 0 \\ u & 1 & 0 \\ w' & v & z \end{pmatrix} \in G_{2m+1},$$

on a

$$\rho(x)\,.\,w = \rho(x)\,.\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ w & 0 & 0 \end{pmatrix} = \begin{pmatrix} y & 0 & 0 \\ u & 1 & 0 \\ w' & v & z \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ w & 0 & 0 \end{pmatrix}\begin{pmatrix} y & 0 & 0 \\ u & 1 & 0 \\ w' & v & z \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ z\,w\,y^{-1} & 0 & 0 \end{pmatrix} = z\,w\,y^{-1}.$$

Le raisonnement s'achève alors exactement comme pour le Théorème 1.

5. Cas où n est impair. Formule de Plancherel. Nous noterons A_{2m+1}' l'ensemble des $x \in G_{2m+1}$ tels que $y = z = 1$, $v = 0$. La formule (18) montre que A_{2m+1}' est un sous-groupe distingué abélien de G_{2m+1}. Soit $\mathfrak{a}_{2m+1}'$ l'algèbre de Lie de A_{2m+1}', c'est-à-dire l'ensemble des matrices

$$\begin{pmatrix} 0 & 0 & 0 \\ u & 0 & 0 \\ w & 0 & 0 \end{pmatrix}.$$

L'application exponentielle de $\mathfrak{a}_{2m+1}'$ sur A_{2m+1}' est un isomorphisme. Donc, pour $e \in E_m$, l'application

$$\begin{pmatrix} 1 & 0 & 0 \\ u & 1 & 0 \\ w & 0 & 1 \end{pmatrix} \to \exp i \operatorname{tr}(e\,w)$$

est un caractère ξ_e de A_{2m+1}'. Nous noterons U_e la représentation unitaire de G_{2m+1} induite par ξ_e.

Tout élément de G_{2m+1} se met de manière unique sous la forme

$$\begin{pmatrix} 1 & 0 & 0 \\ u & 1 & 0 \\ w & 0 & 1 \end{pmatrix}\begin{pmatrix} y & 0 & 0 \\ 0 & 1 & 0 \\ 0 & v & z \end{pmatrix} = \begin{pmatrix} y & 0 & 0 \\ u\,y & 1 & 0 \\ w\,y & v & z \end{pmatrix}.$$

Ainsi, G_{2m+1} est produit semi-direct de A_{2m+1}' et d'un groupe canoniquement isomorphe à $G_m \times G_{m+1}$. En outre

$$(19)\qquad \begin{pmatrix} y' & 0 & 0 \\ 0 & 1 & 0 \\ 0 & v' & z' \end{pmatrix}\begin{pmatrix} y & 0 & 0 \\ u & 1 & 0 \\ w & v & z \end{pmatrix} = \begin{pmatrix} y'\,y & 0 & 0 \\ u & 1 & 0 \\ v'\,u + z'\,w & v' + z'\,v & z'\,z \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ u\,y^{-1}\,y'^{-1} & 1 & 0 \\ (v'\,u + z'\,w)y^{-1}\,y'^{-1} & 0 & 1 \end{pmatrix}\begin{pmatrix} y'\,y & 0 & 0 \\ 0 & 1 & 0 \\ 0 & v' + z'\,v & z'\,z \end{pmatrix}.$$

Par suite, U_e opère dans $L_C{}^2\,(G_m \times G_{m+1})$, $G_m \times G_{m+1}$ étant muni de la mesure de Haar canonique; et, si $(y', z', v') \to f(y', z', v')$ est un élément de $L_C{}^2(G_m \times G_m)$, la formule (19) prouve que, pour

$$x = \begin{pmatrix} y & 0 & 0 \\ u & 1 & 0 \\ w & v & z \end{pmatrix} \in G_{2m+1}$$

on a

$$(20)\quad (U_e(x)f)(y', z', v') = f(y'\,y, z'\,z, v' + z'\,v) \exp i \operatorname{tr}(e(v'\,u + z'\,w)y^{-1}\,y'^{-1}).$$

La formule (20) définit explicitement la représentation U_e.

THÉORÈME 5. *Pour* $e = (\epsilon_{jk}) \in E_m$, *posons* $\epsilon_1 = \epsilon_{m,1}, \ldots, \epsilon_m = \epsilon_{1,m}$.

(i) *La représentation* U_e *admet le caractère infinitésimal* χ_e *défini par*

$$\chi_e(e_{2m+1,1}) = i\epsilon_m, \quad \chi_e\left(\begin{vmatrix} e_{2m,1} & e_{2m,2} \\ e_{2m+1,1} & e_{2m+1,2}\end{vmatrix}\right) = i^2\epsilon_{m-1}\epsilon_m, \ldots,$$

$$\chi_e\left(\begin{vmatrix} e_{m+2,1} & \ldots & e_{m+2,m} \\ \ldots & \ldots & \ldots \\ e_{2m+1,1} & \ldots & e_{2m+1,m}\end{vmatrix}\right) = i^m\epsilon_1\epsilon_2\ldots\epsilon_m.$$

(ii) *Si on se limite aux* $e \in E_m$ *tels que* $\epsilon_1\epsilon_2\ldots\epsilon_m \neq 0$, *les représentations* U_e *sont irréductibles et deux à deux inéquivalentes.*

(iii) *Si* F *est une fonction intégrable sur* G_{2m+1}, *on a*

$$\int_{G_{2m+1}} |F(x)|^2\,dx = (2\pi)^{-m^2-m}\int\ldots\int \operatorname{tr}(U_e(F)^*U_e(F))\,|\epsilon_1\epsilon_2^3\ldots\epsilon_m^{2m-1}|\; d\epsilon_1\,d\epsilon_2\ldots d\epsilon_m.$$

Démonstration. Si

$$x = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ w & 0 & 1\end{pmatrix} \in A_{2m+1},$$

la formule (20) devient

$$(U_e(x)f)(y', z', v') = f(y', z', v')\exp i\operatorname{tr}(e\,z'\,w\,y'^{-1}).$$

La démonstration de (i) s'achève alors exactement comme pour le Théorème 2.

Pour prouver (ii), on raisonne aussi comme pour le Théorème 2. Il s'agit de prouver que, si $\epsilon_1\epsilon_2\ldots\epsilon_m \neq 0$, un élément

$$x = \begin{pmatrix}y & 0 & 0\\ u & 1 & 0\\ w & v & z\end{pmatrix} \in G_{2m+1}$$

qui laisse fixe ξ_e appartient à A_{2m+1}'. Or la condition que x laisse fixe ξ_e se traduit, en vertu de (18), par la condition

$$\operatorname{tr}(e\,w') = \operatorname{tr}(e(v\,u' + z\,w')y^{-1}) \tag{21}$$

quels que soient $w' \in M_m$ et la matrice u' à 1 ligne et m colonnes. Faisant $u' = 0$, ceci impose d'abord $\operatorname{tr}(e\,w') = \operatorname{tr}(y^{-1}\,e\,z\,w')$ quel que soit $w' \in M_m$, donc $e = y^{-1}e\,z$, donc (Lemme 1) $y = z = 1$. La condition (21) se réduit alors à $\operatorname{tr}(e\,v\,u') = 0$ quel que soit u'. Posant

$$v = \begin{pmatrix}\sigma_1\\ \vdots\\ \sigma_m\end{pmatrix} \qquad u' = (\sigma_1' \ldots \sigma_m')$$

on a

$$v\,u' = (\sigma_j\sigma_k'),\ e\,v\,u' = (\epsilon_{j,m-j+1}\,\sigma_{m-j+1}\,\sigma_k'),$$
$$\operatorname{tr}(e\,v\,u') = \sum_j \epsilon_{m-j+1}\,\sigma_{m-j+1}\,\sigma_j',$$

d'où la condition $\epsilon_{m-j+1}\sigma_{m-j+1} = 0$ ($j = 1, \ldots, m$); comme les ϵ_j sont tous $\neq 0$, on en conclut que $v = 0$, d'où $x \in A_{2m+1}'$.

Soit $x \to F(x) = F_1(y, z, v, u, w)$ une fonction intégrable sur G_{2m+1}. Pour $f, f' \in L_C^2(G_m \times G_{m+1})$, on a

$$(22)\quad (U_e(F)f|f') =$$
$$\int\!\!\int\!\!\int\!\!\int\!\!\int\!\!\int\!\!\int\!\!\int F_1(y, z, v, u, w)\, f(y'y, z'z, v' + z'v)\,\overline{f'(y', z', v')}$$
$$\exp i\, \mathrm{tr}(e(v'u + z'w)y^{-1}y'^{-1})\, dy\, dz\, dv\, du\, dw\, dy'\, dz'\, dv'$$
$$= \int \ldots \int F_1(y'^{-1}y, z'^{-1}z, v, u, w)\, f(y, z, v' + z'v)\,\overline{f'(y', z', v')}$$
$$\exp i\, \mathrm{tr}(e(v'u + z'w)y^{-1})\, dy\, dz\, dv\, du\, dw\, dy'\, dz'\, dv'$$
$$= \int \ldots \int F_1(y'^{-1}y, z'^{-1}z, z'^{-1}(v - v'), u, w)\, f(y, z, v)\,\overline{f'(y', z', v')}$$
$$\exp i\, \mathrm{tr}(e(v'u + z'w)y^{-1})\, dy\, dz\, dv\, du\, dw\, dy'\, dz'\, dv'.$$

Donc

$$\mathrm{tr}(U_e(F)^* U_e(F)) = \int \ldots \int \Big| \int F_1(y'^{-1}y, z'^{-1}z, z'^{-1}(v - v'), u, w)$$
$$\exp i\, \mathrm{tr}(e(v'u + z'w)y^{-1})\, du\, dw \Big|^2 dy\, dz\, dv\, dy'\, dz'\, dv'.$$

Comme plus haut, nous identifions l'espace vectoriel M_m à son dual grâce à la forme bilinéaire $(w_1, w_2) \to \mathrm{tr}(w_1 w_2)$. D'autre part, l'espace vectoriel $M_{m,1}$ des matrices u à 1 ligne et m colonnes admet un dual que nous identifions à l'espace vectoriel $M_{1,m}$ des matrices v à m lignes et 1 colonne, en posant

$$\langle u, v \rangle = \mathrm{tr}(v\, u).$$

Alors, la fonction

$$(y, z, v, u, w) \to F_1(y, z, v, u, w)$$

définie sur $G_m \times G_m \times M_{1,m} \times M_{m,1} \times M_m$, admet, par rapport aux variables u et w, une transformée de Fourier, que nous noterons $\bar{F}$, définie sur $G_m \times G_m \times M_{1,m} \times M_{1,m} \times M_m$; et l'on a

$$\mathrm{tr}(U_e(F)^* U_e(F)) = \int \ldots \int |\bar{F}(y'^{-1}y, z'^{-1}z, z'^{-1}(v - v'), y^{-1}ev', y^{-1}ez')|^2$$
$$dy\, dz\, dv\, dy'\, dz'\, dv'$$
$$= \int \ldots \int |\bar{F}(y, z, z'^{-1}(v - v'), y^{-1}y'^{-1}ev', y^{-1}y'^{-1}ez')|^2$$
$$dy\, dz\, dv\, dy'\, dz'\, dv'$$
$$= \int \ldots \int |\bar{F}(y, z, z'^{-1}(v - v'), y'ev', y'ez')|^2\, dy\, dz\, dv\, dy'\, dz'\, dv'$$
$$= \int \ldots \int |\bar{F}(y, z, v, y'ev', y'ez')|^2\, dy\, dz\, dv\, dy'\, dz'\, dv'.$$

Pour $\epsilon_1 \epsilon_2 \ldots \epsilon_m \neq 0$, on a $\det(y'e) = \det e \neq 0$. Donc

$$\mathrm{tr}(U_e(F)^* U_e(F)) = \int \ldots \int |\bar{F}(y, z, v, v', y'ez')|^2 |\det e|^{-1}\, dy\, dz\, dv\, dy'\, dz'\, dv'.$$

Alors, en utilisant le Lemme 3,

$$\begin{aligned}\int \dots \int \operatorname{tr}(U_e(F)^* U_e(F)) |\epsilon_1\epsilon_2^3 \dots \epsilon_m^{2m-1}|\, d\epsilon_1 d\epsilon_2 \dots d\epsilon_m \\
&= \int \dots \int |\bar{F}(y, z, v, v', y'ez')|^2 \epsilon_2^2\epsilon_3^4 \dots \epsilon_m^{2m-2}\, dy\, dz\, dv\, dy'\, dz'\, dv'\, d\epsilon_1\, d\epsilon_2 \dots d\epsilon_m \\
&= \int \dots \int |\bar{F}(y, z, v, v', w)|^2\, dy\, dz\, dv\, dv'\, dw \\
&= (2\pi)^{m^2+m} \int \dots \int |F(y, z, v, u, w)|^2\, dy\, dz\, dv\, du\, dw \\
&= (2\pi)^{m^2+m} \int |F(x)|^2\, dx.\end{aligned}$$

Ceci achève la démonstration du Théorème 5.

6. Cas où n est impair. Caractères globaux des représentations U_e.

THÉORÈME 6. *Soit* $e = (\epsilon_{jk}) \in E_m$; *posons* $\epsilon_1 = \epsilon_{m,1}, \dots, \epsilon_m = \epsilon_{1,m}$; *supposons* $\epsilon_1\epsilon_2 \dots \epsilon_m \neq 0$.

(i) *Soit*

$$x = \begin{pmatrix} y & 0 & 0 \\ u & 1 & 0 \\ w & v & z \end{pmatrix} \to F(x) = F_1(y, z, v, u, w)$$

une fonction indéfiniment dérivable à décroissance rapide sur G_{2m+1}. *Alors, l'opérateur* $U_e(F)$ *est un opérateur à trace.*

(ii) *Il existe une distribution tempérée* T_e *sur* M_m *telle que*

$$\operatorname{tr}(U_e(F)) = \int F_1(1, 1, 0, 0, w)\, dT_e(w).$$

(iii) *Sur l'ensemble* N_m *des* $w \in M_m$ *tels que* $\Delta_2(w)\Delta_3(w) \dots \Delta_m(w) \neq 0$, T_e *coïncide avec la fonction*

$$w \to (2\pi)^{\frac{1}{2}m(m+1)} |\epsilon_1^{-1}\epsilon_2^{-2} \dots \epsilon_m^{-m}| \frac{\exp i\left[\epsilon_1\Delta_m(w) - \epsilon_2 \dfrac{\Delta_{m-1}(w)}{\Delta_m(w)} + \dots + (-1)^{m+1} \epsilon_m \dfrac{\Delta_1(w)}{\Delta_2(w)}\right]}{|\Delta_2(w)\Delta_3(w) \dots \Delta_m(w)|}.$$

Démonstration. Comme dans la démonstration du Théorème 5, introduisons la fonction $(y, z, v, u, w) \to \bar{F}(y, z, v, u, w)$ sur $G_m \times G_m \times M_{1,m} \times M_{1,m} \times M_m$, transformée de Fourier de la fonction $(y, z, v, u, w) \to F_1(y, z, v, u, w)$ par rapport aux variables u et w. La formule (22) prouve que $U_e(F)$ est défini par le noyau

$$K(y, z, v, y', z', v') = \bar{F}(y'^{-1}y, z'^{-1}z, z'^{-1}(v - v'), y^{-1}ev', y^{-1}ez').$$

Comme dans la démonstration du Théorème 3, nous introduisons les notations $||y||$, $||w||$ pour les éléments de G_m, M_m, et nous posons $||v|| = \Sigma_j|\tau_j|$ si $v = (\tau_j) \in M_{1,m}$. On a vu qu'il existe des constantes $A_1 > 0$, $\delta_1 > 0$ telles que

$$||y'^{-1}y|| + ||z'^{-1}z|| + ||y^{-1}ez'|| > A_1(||y|| + ||z|| + ||y'|| + ||z'||)^{\delta_1};$$

d'autre part, il existe des constantes $A_2 > 0$, $\delta_2 > 0$ telles que

$$||v'|| = ||e^{-1}y(y^{-1}ev')|| < A_2(||y|| + ||y^{-1}ev'||)^{\delta_2}$$

et des constantes $A_3 > 0$, $\delta_3 > 0$ telles que

$$||v|| = ||z'(z'^{-1}(v - v')) + v'|| < A_3(||z'|| + ||v'|| + ||z'^{-1}(v - v')||)^{\delta_3}$$

Finalement, il existe des constantes $A > 0$, $\delta > 0$, telles que

$$||y'^{-1}y|| + ||z'^{-1}z|| + ||z'^{-1}(v - v')|| + ||y^{-1}ev'|| + ||y^{-1}ez'|| \\ \geqslant A\,(||y|| + ||y'|| + ||z|| + ||z'|| + ||v|| + ||v'||)^{\delta}.$$

D'après le Lemme 4, K est donc une fonction indéfiniment dérivable à décroissance rapide. Par suite, $U_e(F)$ est un opérateur à trace, et

$$\mathrm{tr}(U_e(F)) = \iiint \bar{F}\,(1, 1, 0, yev, yez)\, dy\, dz\, dv.$$

L'application $(y, z, v) \to (yev, yez)$ de $G_m \times G_m \times M_{1,m}$ dans $M_{1,m} \times M_m$ transforme la mesure $dy\,dz\,dv$ en une distribution tempérée sur $M_{1,m} \times M_m$ d'après le Lemme 4 et les inégalités obtenues plus haut; soit T_e' la transformée de Fourier de cette distribution, qui est aussi une distribution tempérée; alors

$$\mathrm{tr}(U_e(F)) = \int F(1, 1, 0, u, w)\, dT_e'(u, w).$$

On a

$$\bar{F}(1, 1, 0, yev, yez) = \iint F_1(1, 1, 0, u, w) \exp i\, \mathrm{tr}(yevu + yezw)\, du\, dw$$

$$\begin{aligned}
\int \bar{F}\,(1, 1, 0, yev, yez)\, dv
&= \int dv \iint F_1\,(1, 1, 0, u, w) \exp i\, \mathrm{tr}(yevu + yezw)\, du\, dw \\
&= \int |\det e|^{-1}\, dv \iint F_1(1, 1, 0, u, w) \exp i\, \mathrm{tr}(vu + yezw)\, du\, dw \\
&= \int |\det e|^{-1}\, dv \iint F_1(1, 1, 0, u, we^{-1}y^{-1}) \exp i\, \mathrm{tr}(vu + zw) |\det e|^{-m}\, du\, dw \\
&= |\det e|^{-1-m} \int dv \iint F_1(1, 1, 0, u, we^{-1}y^{-1}) (\exp i\, \mathrm{tr}\, w) \\
&\qquad (\exp i\, \mathrm{tr}(vu + z^\circ w))\, du\, dw
\end{aligned}$$

en posant $z = 1 + z^\circ$. D'où

$$\int \bar{F}(1, 1, 0, yev, yez)\, dv = (2\pi)^m |\det e|^{-1-m} \int F_1\,(1, 1, 0, 0, we^{-1}y^{-1}) \\ (\exp i\, \mathrm{tr}\, w)\, (\exp i\, \mathrm{tr}\, z^\circ w)\, dw.$$

Puis, raisonnant comme pour le Théorème 3

$$\iint \bar{F}(1, 1, 0, yev, yez) dz dv = (2\pi)^m |\det e|^{-1-m} (2\pi)^{\frac{1}{2}m(m-1)} \\ \int F_1(1, 1, 0, 0, te^{-1}y^{-1}) (\exp i\, \mathrm{tr}\, t)\, dt$$

où t parcourt l'ensemble des matrices (τ_{jk}) telles que $\tau_{jk} = 0$ pour $j < k$ et où dt est définie par $\Pi_{j \geqslant k} d\tau_{jk}$. Posons $te^{-1} = ze'$, où $z \in G_m$ et $e' = (\epsilon_{jk}') \in E_m$, et $\epsilon'_{m-j+1,j} = \epsilon_j'$. Il vient

$$\mathrm{tr}\, t = \mathrm{tr}(ze'e) = \mathrm{tr}(e'e) = \sum_j \epsilon'_{m-j+1,} \epsilon_j$$

$$\begin{aligned}
dt &= |\epsilon_1 \ldots \epsilon_m| |(\epsilon_1\epsilon_m')^{m-1} (\epsilon_2\epsilon_{m-1}')^{m-2} \ldots (\epsilon_{m-1}\epsilon_2')| d\epsilon_1' \ldots d\epsilon_m'\, dz \\
&= |\epsilon_1^m \epsilon_2^{m-1} \ldots \epsilon_m| |\epsilon_2'^{-1} \epsilon_3'^{-2} \ldots \epsilon_m'^{-m+1}| de' dz.
\end{aligned}$$

D'où

$$\iint \bar{F}(1, 1, 0, yev, yez)\, dz dv$$
$$= (2\pi)^{\frac{1}{2}m(m+1)}|\epsilon_1^{-1}\epsilon_2^{-2} \ldots \epsilon_m^{-m}| \iint F_1(1, 1, 0, 0, ze'y^{-1})\Big(\exp i \sum_j \epsilon'_{m-j+1}\epsilon_j\Big)$$
$$|\epsilon_2'^{j-1} \ldots \epsilon_m'^{j-m+1}|\, de'dz.$$

Supposons maintenant le support de la fonction $w \to F_1\,(1, 1, 0, 0, w)$ compact et contenu dans N_m. On a

$$\iiint \bar{F}(1, 1, 0, yev, yez)\, dy\, dz\, dv$$
$$= (2\pi)^{\frac{1}{2}m(m+1)}|\epsilon_1^{-1}\epsilon_2^{-2} \ldots \epsilon_m^{-m}| \iiint F_1(1, 1, 0, 0, ze'y^{-1})\Big(\exp i \sum_j \epsilon'_{m-j+1}\epsilon_j\Big)$$
$$|\epsilon_2'^{j-1} \ldots \epsilon_m'^{j-m+1}|\, dy\, de'dz$$
$$= (2\pi)^{\frac{1}{2}m(m+1)}|\epsilon_1^{-1} \ldots \epsilon_m^{-m}| \int F_1(1, 1, 0, 0, w)$$
$$\frac{\exp i\left[\epsilon_1\Delta_m(w) + \ldots + (-1)^{m+1}\epsilon_m \dfrac{\Delta_1(w)}{\Delta_2(w)}\right]}{|\Delta_2(w) \ldots \Delta_m(w)|}\, dw.$$

Remarques. 1. Sauf dans le cas $m = 1$ (cas qui est étudié dans **(4)**), on voit que la distribution T_s du théorème n'est pas une mesure.

2. Soit G un groupe de Lie nilpotent simplement connexe. La conjecture suivante paraît vraisemblable : les caractères globaux des représentations unitaires irréductibles de G sont des distributions tempérées sur G (la notion de distribution tempérée se définissant grâce à l'application exponentielle).

Errata à un article antérieur :*

(1) P. 322, 1.4, la conjecture que $\mathfrak{Z}(\mathfrak{g})$ est de type fini est inexacte : on le voit en utilisant le contre-exemple au 14^e problème de Hilbert que vient d'obtenir Nagata.

(2) Les algèbres de Lie définies pp. 322-3 sont, pour les groupes de Lie définis pp. 330-1, les algèbres des champs de vecteurs invariants à droite. Or, dans l'article précédent de cette série, on a toujours utilisé les champs de vecteurs invariants à gauche. Il en résulte qu'il faut changer M_1 en $-M_1$ et M_2 en $-M_2$ dans les formules (9), (12), (15), (18), (21), (24), (27), (30), (33) et dans leurs démonstrations.

(3) Dans la formule (25), les 4 derniers termes du crochet doivent être :

$$-(\lambda^2 + \mu^2)\rho_3\theta - \mu^2\rho_1\rho_2\theta - \tfrac{1}{2}\lambda\mu(\rho_1^2 - \rho_2^2)\theta - \tfrac{1}{2}(\lambda^2 + \mu^2)(\mu\rho_1 - \lambda\rho_2)\theta^2.$$

P. 341, 1.17, le dernier terme doit être : $\frac{1}{2}\lambda\mu(\rho_1{}^2 - \rho_2{}^2)$. Dans la formule (33), il faut :

$$U_\lambda(x_3) = -i\lambda M_2 + \tfrac{1}{2}\, i\lambda M_1^2 \qquad U_\lambda(x_4) = -i\lambda M_1$$

**Sur les représentations unitaires des groupes de Lie nilpotents. III*, Can. J. Math., *10* (1958), pp. 321–48.

Dans la formule (34), il faut

$$\exp i\lambda\left(\rho_5 - \rho_4\theta_1 + \tfrac{1}{2}\,\rho_2^2\theta_1 + \tfrac{1}{2}\,\rho_3\theta_1^2 - \frac{1}{6}\,\rho_2\theta_1^3 - \rho_3\theta_2 + \rho_2\theta_1\theta_2\right).$$

Bibliographie

1. F. Bruhat, *Sur les représentations induites des groupes de Lie*, Bull. Soc. Math. France, *84* (1956), 97–205.
2. J. Dixmier, *Sur les représentations unitaires des groupes de Lie nilpotents, II.*, Bull. Soc. Math. France, *85* (1957), 325–88.
3. I. M. Gelfand and M. A. Neumark, *Unitäre Darstellungen der klassischen Gruppen* (Berlin, 1957).
4. R. Godement, *Mémoire sur la théorie des caractères dans les groupes localement compacts unimodulaires*, J. Math. pures et appl., *30* (1951), 1–110.
5. ———, *Théorie des caractères, II. Définition et propriétés générales des caractères*, Ann. Math., *59* (1954), 63–85.
6. A. Grothendieck, *Produits tensoriels topologiques et espaces nucléaires*, Mem. Amer. Math. Soc., *16* (1955).
7. G. W. Mackey, *Imprimitivity for representations of locally compact groups, I.*, Proc. Nat. Acad. Sc. U.S.A., *35* (1949), 537–45.
8. L. Schwartz, *Théorie des distributions*, II (Paris, Hermann, Act. Sc. et Ind., *1122* [1951]).

Institut Henri Poincaré

NOTE ON GENERALIZED WITT ALGEBRAS

RIMHAK REE

Introduction. Throughout this note K will denote a field of characteristic $p > 0$. Let I be the set $\{1, 2, \ldots, m\}$, and $\mathfrak{G}$ a finite additive group of functions on I with values in K. We assume that $\mathfrak{G}$ is total in the sense that, for any $\lambda_1, \ldots, \lambda_m$ in K, $\sum_{i=1}^m \lambda_i \sigma(i) = 0$ for all σ in G implies all $\lambda_i = 0$. It is clear that $\mathfrak{G}$ is an elementary p-group. Let p^n be the order of $\mathfrak{G}$. A generalized Witt algebra $\mathfrak{L}$ is defined as an algebra over K with basis elements $\{e(\sigma, i) | \sigma \in \mathfrak{G}, i \in I\}$ and the multiplication table

$$e(\sigma, i)e(\tau, j) = \tau(i)e(\sigma + \tau, j) - \sigma(j)e(\sigma + \tau, i). \tag{0.0.1}$$

$\mathfrak{L}$ is a simple Lie algebra except when $p = 2$, $m = 1$.

In the first section of this note we shall prove that the outer derivation algebra of a generalized Witt algebra is abelian, assuming that K is infinite. We shall see that actually a result of Jacobson **(3)** is generalized.

It was shown in **(5)** that any generalized Witt algebra $\mathfrak{L}$ can be reformulated as follows: Let $\mathfrak{A}$ be a commutative associative algebra over K with a unity element, and $D_1, \ldots, D_m$ be derivations of $\mathfrak{A}$ such that:

(1) $[D_i, D_j] = D_iD_j - D_jD_i = 0$ for all i and j;

(2) If $f \in \mathfrak{A}$ and $\lambda_1, \ldots, \lambda_k$ in K are such that $D_if = \lambda_if$ for all i then $f = 0$ or f is a unit in $\mathfrak{A}$;

(3) $\sum_{i=1}^m f_iD_i = 0$, where $f_i \in \mathfrak{A}$, implies $f_i = 0$ for all i.

Now any generalized Witt algebra can be regarded as the subalgebra $\mathfrak{L}(\mathfrak{A}; D_1, \ldots, D_m)$ of the derivation algebra of $\mathfrak{A}$ consisting of all derivations of the form $f_1D_1 + \ldots + f_mD_m$. In the second section of this note we shall consider $\mathfrak{L}(\mathfrak{A}; D_1, \ldots, D_m)$ under the conditions (1) and (2) above only, and extend some results proved in **(5)**.

1. The derivation algebra of a generalized Witt algebra.

We prove the following

THEOREM 1.1. *Let $\mathfrak{L}$ be a generalized Witt algebra over an infinite field K of characteristic $p > 2$. Let $\{e(\sigma, i) | \sigma \in G, i \in I\}$ be a basis of $\mathfrak{L}$. Then any derivation of $\mathfrak{L}$ is the sum of an inner derivation and a derivation δ_1 given by*

$$\delta_1(e(\sigma, i)) = \phi(\sigma)e(\sigma, i) \tag{1.1.1}$$

where ϕ is a linear map of $\mathfrak{G}$ into K.

Received May 12, 1958. This research was supported by the United States Air Force through the Air Research and Development Command under Contract No. AF 49(638)-152.

Proof. First of all we show that we may assume (1.1.2): for any i, $1 \leq i \leq m$, $\sigma(i) = 0$ implies $\sigma = 0$. Suppose (1.1.2) is not satisfied. Since K is infinite and $\mathfrak{G}$ total, we may proceed as in the proof of Lemma 9.1 of (5, p. 533) to obtain an $m \times m$ non-singular matrix (β_{ij}) such that if we define $\sigma[i]$ by

$$\sigma[i] = \sum_{j=1}^{m} \beta_{ij}\,\sigma(j), \qquad (i = 1, \ldots, m),$$

then, for any i, $\sigma[i] = 0$ implies $\sigma = 0$. Define a new basis $\{e[\sigma, i] \mid \sigma \in \mathfrak{G}, i \in I\}$ of $\mathfrak{L}$ by

$$e[\sigma, i] = \sum_{j=1}^{m} \beta_{ij}\, e(\sigma, i).$$

Then by (0.0.1) we have

$$\begin{aligned} e[\sigma, i]\, e[\tau, j] &= \sum_{s,t} \beta_{is}\beta_{jt}\, e(\sigma, s)e(\tau, t) \\ &= \sum_{s,t} \beta_{is}\beta_{jt}\, (\tau(s)e(\sigma + \tau, t) - \sigma(t)e(\sigma + \tau, s)) \\ &= \tau[s]\, e[\sigma + \tau, t] - \sigma[t]\, e[\sigma + \tau, s]. \end{aligned}$$

Thus $\{e[\sigma, i]\}$ satisfies the same multiplication table as $\{e(\sigma, i)\}$ with $\sigma(i)$ replaced by $\sigma[i]$. But here $\sigma[i] = 0$ implies $\sigma = 0$. Suppose that the given derivation is the sum of an inner derivation and a derivation δ_1 given by $\delta_1(e[\sigma, i]) = \phi(\sigma)e[\sigma, i]$, where ϕ is an additive map of $\mathfrak{G}$ into K. Then clearly we have $\delta_1(e(\sigma, i)) = \phi(\sigma)e(\sigma, i)$ also. This shows that we can assume (1.1.2) from the beginning.

Now let δ be the given derivation, and let

$$\delta(e(\sigma, i)) = \sum_{\tau, j} \gamma(\sigma, i; \tau, j)e(\sigma + \tau, j)$$

with coefficients $\gamma(\sigma, i; \tau, j)$ in K. Then from

$$\delta(e(0, 1))e(\sigma, i) + e(0, 1)\delta(e(\sigma, i)) = \sigma(1)\delta(e(\sigma, i))$$

we obtain

$$\gamma(\sigma, i; \tau, j) = \gamma(0, 1; \tau, j)\tau(i)\tau(1)^{-1} \tag{1.1.3}$$

for $i \neq j$ and $\tau \neq 0$, and

$$\sum_{j} \gamma(0, 1; \tau, j)\sigma(j) + \gamma(\sigma, i; \tau, i)\tau(1) = \gamma(0, 1; \tau, i)\tau(i). \tag{1.1.4}$$

By (1.1.3) and (1.1.4) we see easily that

$$\begin{aligned} \delta(e(\sigma, i)) = &\sum_{j} \gamma(\sigma, i; 0, j)e(\sigma, j) \\ &+ e(\sigma, i) \sum_{\tau \neq 0} \sum_{j} \gamma(0, 1; \tau, j)\tau(1)^{-1}e(\tau, j). \end{aligned}$$

Hence δ is the sum of an inner derivation and a derivation δ_1 of the form

$$\delta_1(e(\sigma, i)) = \sum_{j} \gamma(\sigma, i, j)e(\sigma, j) \tag{1.1.5}$$

with coefficients $\gamma(\sigma, i, j)$ in K.

We shall show that $\gamma(\sigma, i, j) = 0$ if $i \neq j$, that $\gamma(\sigma, 1, 1) = \ldots = \gamma(\sigma, m, m)$, and that $\gamma(\sigma, 1, 1)$ is additive with respect to σ. If $m = 1$, then the additivity of $\gamma(\sigma, 1, 1)$ follows immediately from

$$\delta_1(e(\sigma, 1))e(\tau, 1) + e(\sigma, 1)\delta_1(e(\tau, 1)) = \delta_1(e(\sigma, 1)e(\tau, 1)).$$

Hence we shall assume that $m > 1$. Then from

$$\delta_1(e(\sigma, 1))e(\tau, j) + e(\sigma, i)\delta_1(e(\tau, j)) = \delta_1(e(\sigma, i)e(\tau, j))$$

we have, for $i \neq j$,

$$\gamma(\sigma, i, j)\sigma(i) - \gamma(\tau, i, j)\tau(i) = \gamma(\sigma + \tau, i, j)(\sigma(i) - \tau(i)); \tag{1.1.6}$$

$$\sum_k \gamma(\sigma, i, k)\tau(k) = \gamma(\sigma, i, j)\sigma(j) - \gamma(\tau, j, j)\tau(i) + \gamma(\sigma + \tau, j, j)\tau(i) - \gamma(\sigma + \tau, i, j)\sigma(j). \tag{1.1.7}$$

Setting $\sigma = 0$ in (1.1.7) and using the fact that G is total, we have

$$\gamma(0, i, k) = 0 \tag{1.1.8}$$

for all i and k. Set $\tau = -\sigma$, in (1.1.6) and use (1.1.8). Then we have, for any σ and $i \neq j$,

$$\gamma(\sigma, i, j) + \gamma(-\sigma, i, j) = 0. \tag{1.1.9}$$

Replace τ in (1.1.6) by $-\tau$, and use (1.1.9). Then we have

$$\gamma(\sigma, i, j)\sigma(i) - \gamma(\tau, i, j)\tau(i) = \gamma(\sigma - \tau, i, j)(\sigma(i) + \tau(i)).$$

Combining this with (1.1.6) yields

$$\gamma(\sigma - \tau, i, j)(\sigma(i) + \tau(i)) = \gamma(\sigma + \tau, i, j)(\sigma(i) - \tau(i)). \tag{1.1.10}$$

Since $\mathfrak{G}$ is an elementary p-group and $p \neq 2$, $\sigma - \tau$ and $\sigma + \tau$ may be regarded as two arbitrary elements in $\mathfrak{G}$. Hence by (1.1.10) it follows that, for $i \neq j$,

$$\gamma(\sigma, i, j) = \alpha_{ij}\sigma(i), \tag{1.1.11}$$

where α_{ij} are in K and independent of σ. Substituting this in (1.1.7) we obtain

$$\gamma(\sigma, i, i)\tau(i) + \sum_{k \neq i} \alpha_{ik}\sigma(i)\tau(k) = \gamma(\sigma + \tau, j, j)\tau(i) - \gamma(\tau, j, j)\tau(i) - \alpha_{ij}\tau(i)\sigma(j), \tag{1.1.12}$$

which shows that $(\gamma(\sigma + \tau, j, j) - \gamma(\tau, j, j))\tau(i)$ is additive with respect to τ. Hence

$$\gamma(\sigma + \tau, j, j) - \gamma(\tau, j, j) = \gamma(\sigma - \tau, j, j) - \gamma(-\tau, j, j) \tag{1.1.13}$$

for all σ and τ. Let $\sigma = \tau$ in the above and use (1.1.8). Then

$$\gamma(2\tau, j, j) - \gamma(\tau, j, j) = -\gamma(-\tau, j, j). \tag{1.1.14}$$

By (1.1.13) and (1.1.4) we have

$$\gamma(\sigma + \tau, j, j) = \gamma(\sigma - \tau, j, j) + \gamma(2\tau, j, j)$$

which shows that $\gamma(\sigma, j, j)$ is additive with regard to σ, since, as before, $\sigma + \tau$ and $\sigma - \tau$ can be regarded as two arbitrary elements in $\mathfrak{G}$. Now from (1.1.12) we obtain

$$\gamma(\sigma, i, i)\tau(i) + \sum_{k \neq i} \alpha_{ik}\sigma(i)\tau(k) = \gamma(\sigma, j, j)\tau(i) - \alpha_{ij}\tau(i)\sigma(j)$$

for all σ and τ. Using the fact that G is total, we see from the above that $\alpha_{ik} = 0$ for $k \neq i$ and that $\gamma(\sigma, i, i) = \gamma(\sigma, j, j)$ for any i and j. Set $\gamma(\sigma, i, i) = \phi(\sigma)$. Then ϕ is additive, and we have (1.1.1) as desired. Thus Theorem 1.1 is proved.

When is the derivation δ defined by $\delta(e(\sigma, i)) = \phi(\sigma)e(\sigma, i)$, where ϕ is an additive function on G, inner? Let

$$\delta(e(\sigma, i)) = e(\sigma, i) \sum_{\tau, j} \alpha_{\tau, j} e(\tau, j)$$

with $\alpha_{\tau, j} \in K$. Then

$$0 = e(0, i) = \sum_{\tau, j} \alpha_{\tau, j}\tau(i)e(\tau, j).$$

Hence $\tau(i) = 0$, $\tau = 0$, whenever $\alpha_{\tau, j} \neq 0$. From this it follows that δ is inner if, and only if, $\phi(\sigma) = \sum_j \alpha_j \sigma(j)$ with $\alpha_j \in K$. Such additive functions ϕ form clearly an m-dimensional vector space over K. On the other hand, if $\mathfrak{G}$ is an elementary group of order p^n, then all the additive functions on $\mathfrak{G}$ with values in K form an n-dimensional vector space over K. Hence we have

COROLLARY 1.2. *Let $\mathfrak{L}$ be a generalized Witt algebra with basis $\{e(\sigma, i) | \sigma \in \mathfrak{G}, i \in I\}$, where $\mathfrak{G}$ is an elementary p-group of order p^n, and $I = \{1, 2, \ldots, m\}$. Let $\mathfrak{D}$ and $\mathfrak{J}$ be the derivation algebra and the algebra of inner derivations of $\mathfrak{L}$, respectively. Then $\mathfrak{D}/\mathfrak{J}$ is an abelian algebra of dimension $n - m$, provided that the characteristic of K is greater than* 2.

From the above corollary it follows immediately that the number m is uniquely determined by $\mathfrak{L}$. This is, however, proved in **(5**, p. 546**)**. Also, if $m = n$, then every derivation of $\mathfrak{L}$ is inner. This is a result of Jacobson **(3)**.

2. Generalized orthogonal systems. Let $\mathfrak{A}$ be a finite-dimensional commutative associative algebra over the algebraically closed ground field K. We assume that $\mathfrak{A}$ has a unity element.

An ordered set $(D_1, \ldots, D_m)$ of derivations of $\mathfrak{A}$ will be called a *generalized orthogonal* (*g.o.*) *system* if the following conditions (2.1.1.)-(2.1.2) are satisfied:

(2.1.1.) $[D_i, D_j] = D_iD_j - D_jD_i = 0$ *for all i and j;*

(2.1.2) *If $f \in \mathfrak{A}$ and $\lambda_1, \ldots, \lambda_m \in K$ are such that $D_if = \lambda_if$ for all i, then $f = 0$ or f is a unit of $\mathfrak{A}$.*

A g.o. system $(D_1, \ldots, D_m)$ will be called an *o. system* if it satisfies the following condition:

(2.1.3.) *$\sum_{i=1}^m f_i D_i = 0$, where $f_i \in \mathfrak{A}$, implies $f_i = 0$ for all i.*

LEMMA 2.1. *The conditions* (2.1.1.)–(2.1.2) *imply the following:*

(2.1.4) *$D_i f = 0$ for all $i = 1, \ldots, m$ implies $f \in K$.*

Proof. The set $\mathfrak{B}$ of all $f \in \mathfrak{A}$ such that $D_i f = 0$ for all i is clearly a subalgebra of $\mathfrak{A}$, and, moreover, if $0 \neq f \in \mathfrak{B}$ then by (2.1.2) f^{-1} exists and belongs to $\mathfrak{B}$, since $D_i f^{-1} = -f^{-2} D_i f = 0$. Therefore, $\mathfrak{B}$ is a finite extension field of K. Since K is algebraically closed, we have $\mathfrak{B} = K$.

THEOREM 2.2. *For any g.o. system $(D_1, \ldots, D_m)$ there exists a non-void subset $S = \{i_1, \ldots, i_r\}$ of indices $1, \ldots, m$ such that* (2.2.1)–(2.2.2), *below, hold:*

(2.2.1) *$(D_{i_1}, \ldots, D_{i_r})$ is an o. system;*

(2.2.2) *There exists $\alpha_{is} \in K$ such that*

$$D_i = \sum_{s \in S} \alpha_{is} D_s, \qquad (i = 1, \ldots, m).$$

Proof. Let S be a minimal subset of the indices $1, \ldots, m$ with respect to the property: there exist $a_{is} \in \mathfrak{A}$ such that

$$D_i = \sum_{s \in S} a_{is} D_s \qquad (i = 1, \ldots, m). \tag{2.2.3}$$

We may assume without loss of generality that $S = \{1, \ldots, r\}$. Let V be the set of all r-tuples $(f_1, \ldots, f_r)$ of elements $f_s \in \mathfrak{A}$ such that $\sum_s f_s D_s = 0$. Define addition in V componentwise, scalar multiplication by $\alpha(f_1, \ldots, f_r) = (\alpha f_1, \ldots, \alpha f_r)$, $\alpha \in K$. Then V is a finite-dimensional vector space over K. We shall prove (2.1.3) for $(D_1, \ldots, D_r)$ by showing that $V = 0$. Suppose $V \neq 0$. Since $\sum f_s D_s = 0$ implies $\sum_s (D_i f_s) D_s = 0$, the mapping $(f_1, \ldots, f_r) \to (D_i f_1, \ldots, D_i f_r)$ is a linear transformation of V. Since $D_i(D_j f) = D_j(D_i f)$ for all f $\in \mathfrak{A}$, i, and j, and since K is algebraically closed, there exists a non-zero $(f_1, \ldots, f_r) \in V$ and $\lambda_1, \ldots, \lambda_m \in K$ such that

$$(D_i f_1, \ldots, D_i f_r) = \lambda_i (f_1, \ldots, f_r)$$

for $i = 1, \ldots, m$. Then $D_i f_s = \lambda_i f_s$ for all i and s. Then from (2.1.2) it follows that f_s is either 0 or a unit in $\mathfrak{A}$. Since not all f_s are zero, we may assume $f_1 \neq 0$; f_1 is a unit. Then $D_1 = -f_1^{-1} f_2 D_2 - \ldots - f_1^{-1} f_r D_r$. Then every D_i can be written as a linear combination of $D_2, \ldots, D_r$ with coefficients in $\mathfrak{A}$. This contradicts the minimality of S. Thus $V = 0$, and hence (2.1.3) is proved for $(D_1, \ldots, D_r)$.

Now, from (2.1.1) and (2.2.3) it follows that $\sum_s (D_k a_{is}) D_s = 0$ for all $i, k = 1, \ldots, m$. Therefore by (2.2.1), we have $D_k a_{is} = 0$ and hence, by Lemma 2.1, $a_{is} = \alpha_{is} \in K$ for all i and s. This proves (2.2.2).

In order to show that $(D_1, \ldots, D_r)$ is an *o.* system, it remains to be shown that $D_s f = \lambda_s f$, $\lambda_s \in K$, for $s = 1, \ldots, r$ implies that $f = 0$ or f is a unit. This, however, follows easily from (2.2.2) and (2.1.2). Thus the proof of Theorem 2.2 is complete.

COROLLARY 2.3. *A g.o. system $(D_1, \ldots, D_m)$ is an o. system if, and only if, $D_1, \ldots, D_m$ are linearly independent over K.*

COROLLARY 2.4. *If there exists a g.o. system of derivations of $\mathfrak{A}$, then $\mathfrak{A}$ is isomorphic to the group algebra over K of an abelian p-group of type $(p, p, \ldots, p)$.*

Proof. By Theorem 2.2, there exists an *o.* system of derivations of $\mathfrak{A}$. Then Corollary 2.3 follows from Lemma 2.1 above and Theorem 6.10 of **(5)**.

COROLLARY 2.5. *The conditions* (2.1.1)–(2.1.2) *imply the following: If $f, a_1, \ldots, a_m \in \mathfrak{A}$ are such that $D_i f = a_i f$ for all i, then $f = 0$ or f is a unit in $\mathfrak{A}$.*

Proof. Corollary 2.5 follows immediately from Theorem 2.2 above, and Lemma 6.3 of **(5)**.

The following theorem, which also follows immediately from Theorem 2.2, above, and Theorem 6.10 of **(5)**, is a partial generalization of Theorem 6.10 of **(5)**.

THEOREM 2.6. *If $(D_1, \ldots, D_m)$ is a g.o. system, then the subalgebra of the derivation algebra of $\mathfrak{A}$, consisting of all derivations of the form $\sum f_i D_i$, where $f_i \in \mathfrak{A}$, is isomorphic to a generalized Witt algebra.*

Now let $(D_0, \ldots, D_m)$ be a set of derivations of $\mathfrak{A}$, satisfying (2.1.1), and let $a_0, \ldots, a_m \in \mathfrak{A}$ be such that $D_i a_j = D_j a_i$ for all i and j. Then the set $\mathfrak{L} = \mathfrak{L}(D_0, \ldots, D_m; a_0, \ldots, a_m)$ of all derivations of the form $\sum f_i D_i$, where $f_i \in \mathfrak{A}$ satisfy $\sum_i (D_i f_i - a_i f_i) = 0$, forms a subalgebra of the derivation algebra of $\mathfrak{A}$. A special case of such algebras was considered for the first time by Frank **(2)**, and another by Albert and Frank **(1)**. The general case where $(D_0, \ldots, D_m)$ is an *o.* system was considered by Jennings and Ree **(4)**. Here we consider the case where $(D_0, \ldots, D_m)$ is an arbitrary g.o. system.

THEOREM 2.7. *If $(D_0, \ldots, D_m)$ is a g.o. system, then the algebra $L(D_0, \ldots, D_m; a_0, \ldots, a_m)$ is isomorphic either to a generalized Witt algebra or to an algebra of the form $L(D_0', \ldots, D_r'; a_0', \ldots, a_r')$, where $(D_0', \ldots, D_r')$ is an o. system.*

Proof. If $m = 0$, then $(D_0, \ldots, D_m)$ is an o. system, and so our theorem is clear. We shall proceed by induction on m. Assume that Theorem 2.7 is true for $m - 1$. If $(D_0, \ldots, D_m)$ is an *o.* system then our theorem is clear. If $(D_0, \ldots, D_m)$ is not an *o.* system, then, by Theorem 2.2, we may assume without loss of generality that $D_m = \alpha_0 D_0 + \ldots + \alpha_{m-1} D_{m-1}$ with $\alpha_i \in K$. We have

$$D_k\left(a_m - \sum_{i=0}^{m-1} \alpha_i a_i\right) = D_m a_k - \sum_{i=0}^{m-1} \alpha_i D_i a_k = 0$$

for $k = 0, 1, \ldots, m$. Hence

$$a_m - \sum_{i=0}^{m-1} \alpha_i a_i = \alpha$$

belongs to K by Lemma 2.1.

If $\alpha = 0$ then $\mathfrak{L} = \mathfrak{L}(D_0, \ldots, D_m; a_0, \ldots, a_m)$ and $\mathfrak{L}_1 = \mathfrak{L}(D_0, \ldots, D_{m-1}; a_0, \ldots, a_{m-1})$ coincide. This is seen as follows: Let $\sum_0^m f_i D_i \in L$. Then by definition, $\sum_0^m (D_i f_i - a_i f_i) = 0$, and hence

$$\sum_{i=0}^{m-1} (D_i(f_i + \alpha_i f_m) - a_i(f_i + \alpha_i f_m)) = 0.$$

On the other hand,

$$\sum_{i=0}^{m} f_i D_i = \sum_{i=0}^{m-1} (f_i + \alpha_i f_m) D_i.$$

Therefore, $\sum_0^m f_i D_i \in \mathfrak{L}_1$ and hence $\mathfrak{L} < \mathfrak{L}_1$ is proved. Since $\mathfrak{L}_1 < \mathfrak{L}$ is clear, we have $\mathfrak{L} = \mathfrak{L}_1$.

If $\alpha \neq 0$ then $\mathfrak{L} = \mathfrak{L}(D_0, \ldots, D_m; a_0, \ldots, a_m)$ coincides with the set $\mathfrak{L}_2$ of all derivations of the form $\sum_0^{m-1} g_i D_i$, where g_i runs over $\mathfrak{A}$. This is seen as follows: Clearly we have $L < L_2$. Now, for an arbitrary element $\sum_0^{m-1} g_i D_i$ in $\mathfrak{L}_2$, define $f_0, f_1, \ldots, f_m$ by the formulae:

$$f_m = \alpha^{-1} \sum_{i=0}^{m-1} (D_i g_i - a_i g_i);$$

$$f_i = g_i - \alpha_i f_m, \qquad (0 < i < m).$$

Then it is easily seen that $\sum_0^{m-1} g_i D_i = \sum_0^m f_i D_i$, and that

$$\sum_{i=0}^{m} (D_i f_i - a_i f_i) = 0.$$

Therefore $\sum_0^{m-1} g_i D_i \in \mathfrak{L}$, and hence $\mathfrak{L}_2 < \mathfrak{L}$ is proved. Thus we have $\mathfrak{L} = \mathfrak{L}_2$. Since $\mathfrak{L}_2$ is a generalized Witt algebra, this completes the proof of Theorem 2.7.

Consider now a set of derivations $(D_1, \ldots, D_m)$ of $\mathfrak{A}$ satisfying only the condition (2.1.1) and denote by $\mathfrak{L}$ the subalgebra of the derivation algebra of $\mathfrak{A}$ consisting of all derivations of the form $f_i D_i$, where $f_i \in \mathfrak{A}$. Let $\mathfrak{R}$ be the radical of $\mathfrak{A}$, and let $\mathfrak{O}$ be the set of all $f \in \mathfrak{R}$ such that $D_k(D_j(\ldots(D_i f)\ldots))$ $\in \mathfrak{R}$ for any $i, j, \ldots, k$ (the number of indices $i, j, \ldots, k$ is arbitrary). It is easily seen that $\mathfrak{O}$ is an ideal of $\mathfrak{A}$ and that $f \in \mathfrak{O}$ implies $D_i f \in \mathfrak{O}$ for all i. Therefore every D_i induces a derivation $\bar{D}_i$ of the algebra $\bar{\mathfrak{A}} = \mathfrak{A}/\mathfrak{O}$. Since $[\bar{D}_i, \bar{D}_j] = 0$ follows from $[D_i, D_j] = 0$, we can consider the subalgebra $\mathfrak{L}$ of the derivation algebra of $\bar{\mathfrak{A}}$ consisting of all derivations of the form $\sum \bar{f}_i \bar{D}_i$, where $\bar{f}_i \in \bar{\mathfrak{A}}$. Denote by $\bar{f}$ the image of $f \in \mathfrak{A}$ under the natural homomorphism: $\mathfrak{A} \to \bar{\mathfrak{A}}$. Since $\sum f_i D_i = 0$ implies $\sum \bar{f}_i \bar{D}_i = 0$, a mapping ϕ is uniquely

defined by $\phi(\sum f_i D_i) = \sum \bar{f}_i \bar{D}i$. It is easily seen that ϕ is a homomorphism of $\mathfrak{L}$ onto $\bar{\mathfrak{L}}$. The kernel $\mathfrak{J}$ of ϕ consists of elements $\sum f_i D_i$ such that $\sum \bar{f}_i \bar{D}_i = 0$. Note that $\sum \bar{f}_i \bar{D}_i = 0$ if and only if $\sum f_i(D_i g) \in \mathfrak{D}$ for all $g \in \mathfrak{A}$. From this it follows immediately that the ideal $[\mathfrak{J}, \mathfrak{J}]$ of $\mathfrak{L}$ is contained in the algebra $\mathfrak{L}_1$ consisting of all derivations of the form $\sum f_i D_i$, where $f_i \in \mathfrak{D}$. For a positive integer k, denote by $\mathfrak{L}_k$ the algebra of all derivations of the form $\sum f_i D_i$, where $f_i \in \mathfrak{D}^k$. It is easily seen that $[\mathfrak{L}_k, \mathfrak{L}_1] < \mathfrak{L}_{k+1}$ for any k. Since $\mathfrak{D} < \mathfrak{N}$, it follows that $\mathfrak{D}$ is nilpotent, say, $\mathfrak{D}^t = 0$. Then $\mathfrak{L}_t = 0$, and hence $\mathfrak{L}_1$ is nilpotent, and $\mathfrak{J}$ is solvable.

Consider now the algebra $\bar{\mathfrak{L}}$, assuming that every non-unit element in $\mathfrak{A}$ is contained in the radical $\mathfrak{N}$. We shall prove that $(\bar{D}_1, \ldots, \bar{D}_m)$ is a g.o. system of $\bar{\mathfrak{A}}$. Suppose that $\bar{D}_i \bar{f} = \lambda_i \bar{f}$ for all i, and that $\bar{f}$ is a non-unit in $\bar{\mathfrak{A}}$. Then $D_i f = \lambda_i f + g_i$, where $g_i \in \mathfrak{D}$. Since $\bar{f}$ is not a unit f is also not a unit, and hence by our assumption $f \in \mathfrak{N}$. Then from $D_i f - \lambda_i f \in \mathfrak{D}$ it follows easily that $f \in \mathfrak{D}$. Therefore $\bar{f} = 0$, and hence $(\bar{D}_1, \ldots, \bar{D}_m)$ is proved to be a g.o. system. Then, by Theorem 1.6, $\bar{\mathfrak{L}}$ is isomorphic to a generalized Witt algebra.

An associative algebra $\mathfrak{A}$ is called *completely primary* if the set of non-unit elements coincide with the radical of $\mathfrak{A}$. Summarizing the above, we have

THEOREM 2.8. *Suppose that the commutative associative algebra $\mathfrak{A}$ is completely primary. Then for any set of derivatives $(D_1, \ldots, D_m)$ of $\mathfrak{A}$, which satisfies the condition* (2.1.1.), *the algebra $\mathfrak{L}$ consisting of all derivations of the form $\sum f_i D_i$, where $f_i \in \mathfrak{A}$, has a solvable ideal $\mathfrak{J}$ such that $\mathfrak{L}/\mathfrak{J}$ is isomorphic to a generalized Witt algebra.*

Similarly we may obtain the following

THEOREM 2.9. *Suppose that the commutative associative algebra $\mathfrak{A}$ is completely primary. Then for any set of derivations $(D_1, \ldots, D_m)$ of $\mathfrak{A}$, which satisfies the condition* (2.1.1), *an algebra $\mathfrak{L}$ of the form $\mathfrak{L}(D_0, \ldots, D_m; a_0, \ldots, a_m)$ has a solvable ideal $\mathfrak{J}$ such that $\mathfrak{L}/\mathfrak{J}$ is isomorphic either to a generalized Witt algebra or to an algebra of the form $\mathfrak{L}(E_0, \ldots, E_r; b_0, \ldots, b_r)$, where $(E_0, \ldots, E_r)$ is an o. system of derivations of the group algebra over K of an abelian group of type $(p, \ldots, p)$.*

REFERENCES

1. A. A. Albert and M. S. Frank, *Simple Lie algebras of characteristic p*, Rendiconti del Sem. Mat., Univ. e Politech di Torino, *14* (1955), 117–39.
2. M. S. Frank, *A new class of simple Lie algebras*, Proc. Nat. Acad. Sci., *40* (1954), 713–18.
3. N. Jacobson, *Abstract derivation and Lie algebras*, Trans. Amer. Math. Soc., *42* (1937), 206–24.
4. S. A. Jennings and Rimhak Ree, *On a class of Lie algebras of characteristic p*, Trans. Amer. Math. Soc., *84* (1957), 192–207.
5. Rimhak Ree, *On generalized Witt algebras*, Trans. Amer. Math. Soc., *83* (1956), 510–46.

The University of British Columbia

SUPERSOLUBLE IMMERSION

REINHOLD BAER

Supersoluble immersion of a normal subgroup K of a finite group G shall be defined by the following property:

If σ is a homomorphism of G, and if the minimal normal subgroup J of G^σ is part of K^σ, then J is cyclic (of order a prime).

Our principal aim in the present investigation is the proof of the equivalence of the following three properties of the normal subgroup K of the finite group G:

(i) K is supersolubly immersed in G.

(ii) $K/\phi K$ is supersolubly immersed in $G/\phi K$.

(iii) If θ is the group of automorphisms induced in the p-subgroup U of K by elements in the normalizer of U in G, then $\theta' \theta^{p-1}$ is a p-subgroup of θ.

Though most of our discussion is concerned with the proof of this theorem, some of our concepts and results are of independent interest. In § 1 we investigate groups G such that $G' G^{p-1}$ is a p-group. In § 2 some new and useful characterizations of supersoluble groups are obtained. In § 3 we substitute for supersoluble immersion the concept of a supersoluble pair which consists of a group G and a group θ of automorphisms of G meeting the following requirement:

If L is a θ-admissible normal subgroup of G, then every minimal θ-admissible normal subgroup of G/L is cyclic (of order a prime).

These supersoluble pairs are somewhat easier to handle than supersoluble immersion, though their investigation is, for all practical purposes, equivalent to that of supersoluble immersion.

Notations

G' = commutator subgroup of G.

ZG = centre of G.

ϕG = Frattini subgroup of G = intersection of all the maximal subgroups of G.

p-elements and p-groups are elements and groups of order a power of the prime p.

G^k = subgroup of G, generated by all the kth powers of elements in G.

G is a group of exponent e, if $G^e = 1$.

G is p-closed, if products of p-elements are p-elements.

If U is a subgroup of G, then NU is the normalizer and CU the centralizer of U in G.

Received October 14, 1958.

θ is an irreducible group of automorphisms of the group G, if 1 and G are the only θ-admissible subgroups of G.

All groups considered are *finite*.

0. We begin with a survey of the salient facts of the theory of finite supersoluble groups; for details cf. (**1**; **2**, § 11) and (**5**). A group G is termed *supersoluble*, if every epimorphic image, not 1, of G possesses a cyclic normal subgroup different from 1. This implies the apparently stronger fact that the minimal normal subgroups of the epimorphic images of supersoluble groups are cyclic of order a prime. Subgroups, epimorphic images, and direct products of supersoluble groups are likewise supersoluble. Extensions of supersoluble groups by supersoluble groups are, in general, not supersoluble; but extensions of cyclic groups by supersoluble groups and central extensions of supersoluble groups by supersoluble groups are supersoluble.

HUPPERT'S THEOREM: *The following three properties of G are equivalent: G is supersoluble; $G/\phi G$ is supersoluble; every maximal subgroup of G has index a prime.*

If G is supersoluble, then its commutator subgroup G' is nilpotent; and G has the

Sylow Tower Property of supersoluble groups: If H is an epimorphic image of G and p is a maximal prime divisor of the order of H, then the totality P of p-elements in H is a characteristic p-subgroup of H; in other words: H is p-closed.

1. In this section we are going to discuss a very special class of supersoluble groups which, however, will prove important in the sequel.

We recall that it is customary to term exponent of a group G the l.c.m. of the orders of the elements in G. For our purpose it will be more convenient to say that *G is a group of exponent e whenever $G^e = 1$*, that is, whenever e is some common multiple of the orders of the elements in G. If p is a prime, then the group G is termed *p-closed*, whenever products of elements of order a power of p are again elements of order a power of p. This is equivalent to requiring the existence of one and only one p-Sylow subgroup which is then a characteristic p-subgroup of index prime to p; and this characteristic p-subgroup of G shall be termed the p-component of G.

Definition. If the group G is p-closed, and if G/P is abelian of exponent $p - 1$, where P is the p-component of G, then G is strictly p-closed.

If G is strictly p-closed, then its commutator subgroup G' and the subgroup G^{p-1} generated by the $(p - 1)$th powers of elements in G are characteristic p-subgroups. If conversely G' and G^{p-1} are p-subgroups, then $G' G^{p-1}$ is a characteristic p-subgroup of G such that $G/G'G^{p-1}$ is abelian of exponent $p - 1$. Hence *G is strictly p-closed if, and only if, G' and G^{p-1} are p-subgroups of G.*

It is easy to verify that subgroups, epimorphic images, and direct products of strictly p-closed groups are again strictly p-closed. Likewise, extensions of p-groups by strictly p-closed groups are strictly p-closed.

Consider a strictly p-closed group G. Suppose that M is a minimal normal subgroup of an epimorphic image H of G. Then H is likewise strictly p-closed. Hence $H'H^{p-1}$ is a characteristic p-subgroup of H and at the same time the p-Sylow subgroup of H. If the order of M is prime to p, then

$$[M, H'H^{p-1}] \leqslant M \cap H'H^{p-1} = 1;$$

and if the order of M is divisible by p, then M is part of the p-component $H'H^{p-1}$ of H. Application of a well-known property of p-groups shows that in this case $M \cap Z(H'H^{p-1}) \neq 1$; and this implies $M \leqslant Z(H'H^{p-1})$ because of the minimality of M. Thus we have shown again that $[M, H'H^{p-1}] = 1$. Hence condition (viii) of (**2**, p. 184, Theorem 1) is satisfied by G; and this implies that *strictly p-closed groups are supersoluble.* This important fact has various consequences.

THEOREM 1.1. *G is a cyclic group of order p if, and only if, G is a p-group, not* 1, *possessing an irreducible and strictly p-closed group of automorphisms.*

Proof. If G is cyclic of order p, then its group of automorphisms is cyclic of order $p - 1$, proving the necessity of our condition. If conversely, $G \neq 1$ is a p-group and θ is an irreducible and strictly p-closed group of automorphisms of G, then we recall that G is a normal subgroup of its own holomorph and that we may form consequently the subgroup $G\theta$ of the holomorph of G. Since θ is irreducible, G is a minimal normal subgroup of $G\theta$. Since G is a p-group and θ is strictly p-closed, $G\theta$ is likewise strictly p-closed. Hence $G\theta$ is in particular supersoluble; and this implies that its minimal normal subgroup G is cyclic. Thus G is of order p.

For the convenience of the reader we insert here some well-known facts concerning automorphisms of p-groups.

LEMMA. *The automorphism σ of the p-group G is of order a power of p, if it satisfies one of the following conditions:*

(a) *σ induces the identity automorphism in $G/\phi G$ or*

(b) *there exists a σ-admissible normal chain of G in whose factors the identity automorphism is induced by σ.*

Proof. The sufficiency of condition (a) is contained in a result due to P. Hall (**4**, p. 38). Assume next the existence of σ-admissible subgroups $U(i)$ of G such that

$$1 = U(0),\ U(i) \text{ is a normal subgroup of } U(i+1),\ U(k) = G,$$

$$\sigma \text{ induces the identity automorphism in every } U(i + 1)/U(i).$$

Then σ certainly induces a p-automorphism in $U(0)$. We may therefore make the inductive hypothesis that σ induces a p-automorphism in $U(i)$ for some

$i < k$. There exists consequently a positive integer n such that σ^{p^n} induces the identity automorphism in $U(i)$. Hence σ^{p^n} induces the identity automorphism both in $U(i)$ and in $U(i+1)/U(i)$. It is well known (and may be verified by a simple computation) that σ^{p^n} induces a p-automorphism in $U(i+1)$. Thus we have shown that σ induces a p-automorphism in $U(i+1)$, completing our inductive argument. Hence σ is a p-automorphism of $U(k) = G$.

THEOREM 1.2. *A group G with $ZG = 1$ is strictly p-closed if, and only if, maximal subgroups of G are either normal or else have index p in G.*

Proof. If G is strictly p-closed, then $G'G^{p-1}$ is a p-subgroup of G. If the maximal subgroup S of G is not normal, then in particular G' is not part of S. Hence $G = SG'$. Since G' is a p-subgroup, this implies that $[G:S]$ is a power of p. But strictly p-closed groups are supersoluble; and the maximal subgroups of supersoluble groups have index a prime. Hence $[G:S] = p$, proving the necessity of our condition.

Assume conversely the validity of our condition. Denote by P some p-Sylow subgroup of G and by NP its normalizer in G. If P were not normal, then $NP \neq G$ and there would exist a maximal subgroup S of G containing NP. From $P \leqslant NP \leqslant S$ we conclude that $[G:S]$ is prime to p; and this implies by hypothesis that S is a normal subgroup of G. Consequently S contains every p-Sylow subgroup of G as its own p-Sylow subgroup so that p-Sylow subgroups of G are conjugate in S. Application of the Frattini argument shows that $G = S \cdot NP = S < G$, a contradiction proving the normality of P and the p-closure of G. Since every maximal subgroup of G which contains P has index prime to p, these maximal subgroups are, by hypothesis, normal. Consequently every maximal subgroup of G/P is normal; and this implies by Wielandt's Theorem the nilpotency of G/P; see (**7**, p. 108, Satz 13). Application of Schur's Theorem shows the existence of a complement D to P in G, since $[G:P]$ is prime to p (**7**, p.125, Satz 25). Since $G/P \simeq D$, this subgroup D of G is nilpotent too. Every maximal subgroup of G has, by hypothesis, a prime index. Application of Huppert's Theorem shows the supersolubility of G; see (**5**, p. 416, Satz 9) or (**2**, p. 184, Theorem 1). Consider next normal subgroups A and B of G satisfying $A < B \leqslant P$ and $[B:A] = p$. Since G induces in the cyclic group B/A of order p a cyclic group of automorphisms whose order is a divisor of $p - 1$, it follows in particular that $[B, D'D^{p-1}] \leqslant A$. Since G has been shown to be supersoluble, there exist normal subgroups $A(i)$ of G such that

$$1 = A(0), A(i) < A(i+1), A(k) = P, \quad [A(i+1):A(i)] = p.$$

From what we have shown just now it follows that $[A(i+1), D'D^{p-1}] \leqslant A(i)$ for every i.

In other words: every element in $D'D^{p-1}$ induces an automorphism in P which in turn induces the identity automorphism in every $A(i+1)/A(i)$. By Lemma (b) such an automorphism has order a power of p. Consequently

every element in $D'D^{p-1}$ induces the identity automorphism in P. If $D'D^{p-1}$ were not 1, then we would deduce $D'D^{p-1} \cap ZD \neq 1$ from the nilpotency of D. Elements in $D'D^{p-1} \cap ZD$ commute with every element in P and every element in D; and they belong consequently to the centre of $PD = G$. But $ZG = 1$ by hypothesis and hence $D'D^{p-1} \cap ZD = 1$, a contradiction which proves that $D'D^{p-1} = 1$. It follows from $G/P \simeq D$ that $G'G^{p-1} < P$; and this completes the proof of the strict p-closure of G.

Remark. Note that $ZG = 1$ was not needed for the proof of the necessity of our condition. The example of suitably selected nilpotent groups shows that $ZG = 1$ is indispensable for the proof of the sufficiency of our condition.

THEOREM 1.3. *A group G is strictly p-closed if, and only if,*

(a) ***elements in G do not induce automorphisms of order p in subgroups of order prime to p and***

(b) ***subgroups of order prime to p are abelian of exponent $p - 1$.***

Proof. Assume first the existence of a normal p-subgroup P of G such that G/P is abelian of exponent $p - 1$. Consider a subgroup U of order prime to p. Then $P \cap U = 1$ so that U is isomorphic to the subgroup PU/P of G/P. Since the latter gruop is abelian of exponent $p - 1$, so is U. If furthermore the element g in G induces in U an automorphism of order a power of p, then we may assume without loss of generality that g is a p-element. As such g belongs to P and the commutators $[g, u]$ for u in U belong to $P \cap U = 1$. Hence g commutes with every element in U and so induces the identity automorphism in U. This proves the necessity of (a) and (b).

If strict p-closure were not a consequence of (a) and (b), then there would exist a group G of minimal order satisfying (a), (b), without being strictly p-closed. Every subgroup of G meets requirements (a) and (b). Because of the minimality of G it follows that

(1) every proper subgroup of G is strictly p-closed.

Since G is not strictly p-closed, it is certainly not a p-group. Consequently there exists a prime $q \neq p$ dividing the order of G. Denote by Q a q-Sylow subgroup of G. By (b), $Q \neq 1$ is abelian of exponent $p - 1$. If g is an element in the normalizer NQ of Q, then g is the product $g = g'g''$ of an element g' of order prime to p and an element g'' of order a power of p both of which belong to NQ. Since $Q\{g'\}$ is of order prime to p, it is by (b) abelian so that g' belongs to the centralizer of Q. It is a consequence of (a) that g'' belongs to the centralizer of Q. Thus NQ is the centralizer of the q-Sylow subgroup Q. Hence we may apply Burnside's Theorem asserting the existence of a normal subgroup T of G complementary to Q, (**7**, p. 133, Satz 4). T is a proper subgroup of G, since $Q \neq 1$. Hence T is, by (1), strictly p-closed. Consequently the totality P of p-elements in T is a characteristic p-subgroup of T whose index $[T:P]$ is prime to p. Since P is a characteristic subgroup of the normal subgroup T, P is a normal subgroup of G. Since $[G:T]$ is a power

of q, namely the order of Q, the index $[G:P]$ is prime to p. Application of Schur's Theorem shows the existence of a complement C of P in G. Since $C \simeq G/P$ is of order prime to p, it is by (b) abelian of exponent $p - 1$. Hence G is strictly p-closed, a contradiction proving our theorem.

2. In this section we derive a number of properties of supersoluble groups. Some of them are of independent interest and all of them will be needed in the sequel.

THEOREM 2.1. *The following properties of the group G are equivalent.*

(i) *G is supersoluble.*

(ii) *NU/CU is, for every p-subgroup U of G, strictly p-closed.*

(iii) *The Sylow Tower Property of supersoluble groups is satisfied by G; and NP/CP is, for every p-Sylow subgroup P of G, strictly p-closed.*

Proof. Assume first the supersolubility of G; and consider a p-subgroup U of G. The group θ of automorphisms, induced in U by elements in NU, is essentially the same as NU/CU. Since G is supersoluble, so is its subgroup NU. Since U is a normal p-subgroup of the supersoluble group NU, there exist normal subgroups $U(i)$ of NU such that

$$1 = U(0),\ U(i) < U(i+1),\ [U(i+1):U(i)] = p,\ U(k) = U.$$

Normal subgroups of NU which are part of U are θ-admissible. Thus every $U(i)$ is θ-admissible. Denote by θ^* the totality of those automorphisms in θ which induce the identity automorphism in every $U(i+1)/U(i)$. Clearly θ^* is a normal subgroup of θ. An immediate application of § 1, Lemma (b) shows that every automorphism in θ^* is a p-automorphism. Thus θ^* is a normal p-subgroup of θ. Since every $U(i+1)/U(i)$ is cyclic of order p, its group of automorphisms is cyclic of order $p - 1$. The automorphisms in $\theta^{p-1}\theta'$ induce consequently the identity automorphism in every $U(i+1)/U(i)$. Hence $\theta^{p-1}\theta' \leq \theta^*$. The isomorphic groups θ and NU/CU are therefore strictly p-closed, proving that (ii) is a consequence of (i).

Assume next the validity of (ii). Consider a subgroup S of G and a minimal prime divisor p of the order of S. If U is a p-subgroup of S, then NU/CU is, by (ii), strictly p-closed. It follows that $[NU \cap S]/[CU \cap S]$ is likewise strictly p-closed. But p is a minimal prime divisor of the order of S. Hence $[NU \cap S]/[CU \cap S]$ is a p-group. Thus p-automorphisms only are induced in U by elements in S. Consequently we may apply a result that we derived elsewhere assuring the validity of the Sylow Tower Property of supersoluble groups in G (**3**, Theorem 6.2). It follows that (iii) is a consequence of (ii).

Assume finally the validity of (iii). If K is a normal subgroup of G and p a maximal prime divisor of $[G:K]$, then the totality P^* of p-elements in $G^* = G/K$ is a characteristic p-subgroup of index prime to p. If P is a p-Sylow subgroup of G, then $P^* = KP/K$; and we note that KP is a normal subgroup of G, since P^* is a characteristic subgroup of G^*. Application of the Frattini argument shows therefore $G = (KP)NP = K \cdot NP$; and now one sees without

difficulty that the group of automorphisms induced in P^* by elements in G^* is an epimorphic image of the group of automorphisms induced in P by elements in NP. The latter group is essentially the same as NP/CP. Since P is a p-Sylow subgroup of G, we deduce strict p-closure of NP/CP from (iii). Consequently a strictly p-closed group θ of automorphisms is induced in P^* by elements in G^*. There exists a minimal normal subgroup M^* of G^* which is part of P^*. The group θ^* of automorphisms which are induced in M^* by elements in G^* is an epimorphic image of θ. Hence θ^* is strictly p-closed and, because of the minimality of M^*, irreducible. Application of Theorem 1.1 shows that M^* is cyclic of order p. Hence G is supersoluble so that (i) is a consequence of (iii), q.e.d.

Remark 2.1. It is impossible to omit the first half of condition (iii) as may be seen from the following example. Assume that p and q are primes and that q is a divisor of $p - 1$. Then there exists a group A of order pq which possesses a normal subgroup B of order p such that the elements in A induce in B a group of automorphisms or order q. Clearly A is supersoluble, but not cyclic. Next denote by K an elementary abelian q-group of order q^{pq} and let G be an extension of K by A such that A acts as a regular permutation group on a basis of K (we may choose G as a splitting extension of K by A). The group of automorphisms induced in K by elements in G is isomorphic to A and hence not strictly q-closed. This implies in particular that G, though soluble, is not supersoluble (Theorem 2.1). A q-Sylow subgroup of G is an extension of K by a cyclic group of order q. Since A is an extension of a p-group by a q-group and not cyclic, one sees that q-Sylow subgroups of G are their own normalizers. Hence $NQ/CQ = Q/ZQ$ is, for every q-Sylow subgroup Q of G, a q-group. The p-Sylow subgroups of G are cyclic of order p. Their normalizers may contain elements in K; but these would belong to their centralizers. It follows that NP/CP is cyclic of order q for every p-Sylow subgroup P of G; and such a group is strictly p-closed, since q is a divisor of $p - 1$. Thus G is soluble, but not supersoluble; and G satisfies these cond half of condition (iii), but not the Sylow Tower Property of supersoluble groups.

Remark 2.2. Using results derived by us elsewhere (**3**, Theorem 6.2) one shows the equivalence of the three conditions of Theorem 2.1 with the following property:

If p is a minimal prime divisor of the order of the subgroup S of G, then S is completely p-normal; and NP/CP is, for every p-Sylow subgroup P of G, strictly p-closed.

THEOREM 2.2. *Assume that P is a p-Sylow subgroup of a supersoluble group G.*

(a) *$(NP)'(NP)^{p-1}$ is the direct product of a p-group and a group of order prime to p; and $G = NP$ in case p is the maximal prime divisor of the order of G.*

(b) *$P \cap \phi G < \phi P$; and $P \cap \phi G = \phi P$ in case p is the maximal prime divisor of the order of G.*

Proof. We note first that NP/CP is essentially the same as the group of automorphisms induced in P by elements in NP. It is a consequence of Theorem 2.1 that this group of automorphisms is strictly p-closed. Consequently $(CP)(NP)'(NP)^{p-1}/CP$ is a p-group. Since the p-Sylow subgroup P of G is a normal subgroup of NP, this implies

$$(NP)'(NP)^{p-1} < P \cdot CP.$$

Since P is also a normal subgroup of $P \cdot CP$ whose index is prime to p, it follows that $P \cap CP = ZP$ is a normal p-subgroup of CP whose index in CP is prime to p. By Schur's Theorem there exists a complement Q of $P \cap CP$ in CP; see, for instance (**7**, p. 125, Satz 25). Hence

$$P \cdot CP = P[P \cap CP]Q = PQ.$$

Since Q is part of the centralizer of P, $P \cdot CP$ is the direct product of P and Q. Since the elements in Q are of order prime to p, it follows that $P \cdot CP$ is the direct product of a p-group and a group of order prime to p. But this property is subgroup inherited. Hence $(NP)'(NP)^{p-1}$ is the direct product of a p-group and a group of order prime to p. That $G = NP$ in case p is the maximal prime divisor of the order of G, is a consequence of Theorem 2.1 (the Sylow Tower Property of supersoluble groups).

Denote by A the set of all those elements in G whose orders are divisible by primes greater than p only. Because of the Sylow Tower Property of supersoluble groups A is a characteristic subgroup of G whose order is divisible by primes greater than p only. The product AP is likewise a characteristic subgroup of G. It consists of just those elements in G whose orders are not divisible by primes smaller than p. By Schur's Theorem or by P. Hall's characteristic property for soluble groups there exists a complement B of A in G, since $o(A)$ and $[G:A]$ are relatively prime. Since B is isomorphic to G/A, the complement B contains a p-Sylow subgroup of G; and since any two p-Sylow subgroups of G are conjugate in G, we may assume without loss in generality that $P < B$. Since AP is a characteristic subgroup of G, $P = AP \cap B$ is a normal subgroup of B. The characteristic subgroup ϕP of P is consequently a normal subgroup of B. Let $B^* = B/\phi P$ and $P^* = P/\phi P$. Then P^* is an elementary abelian p-group, the p-Sylow subgroup of B^* and characteristic in B^*. Since B^* is supersoluble, the elements in B^* induce in P^* a strictly p-closed group θ of automorphisms. This group θ is essentially the same as B^*/CP^*. Since P^* is abelian, $P^* < CP^*$ so that $[B^*:CP^*]$ is prime to p. Hence θ is a strictly p-closed group of order prime to p; in other words θ is abelian of exponent $p - 1$. Since the group θ of order prime to p acts on the elementary abelian p-group P^*, it is completely reducible (Maschke's Theorem) (**6**, p. 81, Theorem 46). This signifies that every θ-admissible

subgroup of P^* possesses in P^* a θ-admissible complement. Because of the supersolubility of B^* minimal normal subgroups of B^* which are contained in P^* have order p. It follows that P^* is the direct product of θ-admissible cyclic groups of order p. This in turn implies that the intersection of all maximal θ-admissible subgroups of P^* is equal to 1. Consider now some maximal θ-admissible subgroup M of P^*. Then M has index p in P^* and is a normal subgroup of B^*. There exists, by Schur's Theorem, a complement D of P^* in B^*. It is clear that MD is a maximal subgroup of B^*. It follows now that

$$P^* \cap \phi B^* = 1.$$

Every maximal subgroup of B^* has the form $S/\phi P$ with S a maximal subgroup of B. If J is the intersection of all these maximal subgroups S, then we deduce $P \cap J = \phi P$ from $P^* \cap \phi B^* = 1$. If S is a maximal subgroup of B, then AS is a maximal subgroup of G, since B is a complement of the characteristic subgroup A of G. From $P \cap J = \phi P$ we deduce now that $P \cap \phi G \leqslant \phi P$.

Suppose finally in particular that p is the maximal prime divisor of the order of G. Then P is a characteristic subgroup of G (Sylow Tower Property of supersoluble groups). Consequently ϕP is a characteristic subgroup of G too. We recall that maximal subgroups of supersoluble groups have index a prime. If the maximal subgroup S of G does not contain P, then consequently $[G:S] = p$. From $G = PS$ we deduce now that $[P:P \cap S] = p$, since P is a characteristic subgroup of G. It follows that $\phi P \leqslant P \cap S \leqslant S$. Thus we have shown $\phi P \leqslant \phi G$. Hence

$$P \cap \phi G \leqslant \phi P \leqslant P \cap \phi G,$$

proving $\phi P = P \cap \phi G$ in case p is the maximal prime divisor of the order of G.

Remark. Consider primes p, q such that q^2 is a divisor of $p - 1$. Then there exists an extension G of a cyclic group P of order p by a cyclic group of order q^2 such that the elements in G induce in P a cyclic group of automorphisms whose order is q^2. Every q-Sylow subgroup of G is cyclic of order q^2 and a maximal subgroup of G; and there exist two q-Sylow subgroups of G with intersection 1. Hence $\phi G = 1$. If Q is a q-Sylow subgroup of G, then ϕQ is cyclic of order q. Hence

$$Q \cap \phi G = 1 < \phi Q$$

showing the impossibility of improving (b).

3. We are now ready to turn to the study of supersoluble immersion.

Definition. A normal subgroup K of G is supersolubly immersed in G if to every homomorphism σ of G with $K^\sigma \neq 1$ there exists a cyclic normal subgroup $A \neq 1$ of G^σ such that $A \leqslant K^\sigma$.

This implies the apparently stronger property mentioned in the introduction: If K is a supersolubly immersed normal subgroup of G, if σ is a homomorphism of G, and if a minimal normal subgroup M of G^{σ} is part of K^{σ}, then M is cyclic of order a prime.

We note that S is a supersoluble subgroup of G if S contains the supersolubly immersed normal subgroup K of G, and if S/K is supersoluble. This important property has two interesting consequences:

(a) The product of all supersolubly immersed normal subgroups of G is a supersolubly immersed characteristic subgroup of G.

(b) Every maximal supersoluble subgroup of G contains every supersolubly immersed normal subgroup of G.

Examples show, however, that the intersection of all maximal supersoluble subgroups of G may actually be greater than the product of all supersolubly immersed normal subgroups of G.

Much use will be made of the following theorem: If the normal subgroup K of G is supersolubly immersed in G, then the elements in G induce in K a supersoluble group of automorphisms (**5**, p. 420, Satz 12).

This leads us to the following companion concept.

Definition. A group G and a group θ of automorphisms form a supersoluble pair if the minimal θ-admissible normal subgroups of G/T, for T a θ-admissible normal subgroup of G, are cyclic (of order a prime).

If, for instance, T is a supersolubly immersed normal subgroup of the group G and Γ is the group of automorphisms, induced in T by elements in G, then T, Γ is clearly a supersoluble pair. In a way the converse is true too. For consider a supersoluble pair G, θ. Then let H be the holomorph of G. This contains G as a normal subgroup and it also contains θ. Their product $G\theta$ is an extension of G by θ which realizes in G in an obvious way the automorphism group θ. We shall refer to this splitting extension of G by θ as to *the product of the group G and its group θ of automorphisms.* Since G, θ is a supersoluble pair, it is quite obvious that G is supersolubly immersed in $G\theta$. It follows in particular that the group Γ of automorphisms induced in G by $G\theta$ is supersoluble. Since $\theta < \Gamma$, we have shown the supersolubility of θ. Incidentally we have shown that G, Γ is a supersoluble pair where the group Γ of automorphisms of G is the compositum of θ and the group of inner automorphisms of G.

We may summarize the principal results of the preceding discussion as follows:

The following properties of the normal subgroup T of the group G are equivalent.

(i) *T is supersolubly immersed in G.*

(ii) *A supersoluble pair is formed by T and the group Γ of automorphisms induced in T by G.*

(iii) *$T\Gamma$ is supersoluble.*

We mention finally the following easily verified inheritance properties: If G, θ is a supersoluble pair, and if U is a θ-admissible subgroup of G, then (if we denote the group of automorphisms induced in U by θ likewise by θ) U, θ is a supersoluble pair. If J is a θ-admissible normal subgroup of G, then (if we denote the group of automorphisms induced in G/J by θ likewise by θ) $G/J, \theta$ is a supersoluble pair. If, furthermore, θ is a group of automorphisms of the group G, if L is a cyclic θ-admissible normal subgroup, and if $G/L, \theta$ is a supersoluble pair, then G, θ is a supersoluble pair.

If X is a subgroup of G, it will be convenient to denote by θ_X the totality (subgroup) of X-preserving automorphisms in the group θ.

THEOREM 3.1. *If a group θ of automorphisms of the group G contains all the inner automorphisms of G, then the following properties of the pair G, θ are equivalent:*

(i) *G, θ is a supersoluble pair.*

(ii) *If U is a p-subgroup of G, then θ_U induces a strictly p-closed group of automorphisms in U.*

(iii) *G has the Sylow Tower Property of supersoluble groups; and if P is a p-Sylow subgroup of G, then θ_P induces a strictly p-closed group of automorphisms in P.*

(iv) *θ has the Sylow Tower Property of supersoluble groups; and if Σ is a subgroup of θ and P a Σ-admissible p-Sylow subgroup of G, then maximal Σ-admissible subgroups of P have index p in P.*

Proof. If G, θ is a supersoluble pair, then their product $G\theta$ is a supersoluble group. If U is a p-subgroup of G, then the normalizer of U in $G\theta$ induces in U a strictly p-closed group Γ of automorphisms (Theorem 2.1). Since θ_U is by definition part of the normalizer of U in $G\theta$, it follows that θ_U induces in U a subgroup of Γ. Since the latter is strictly p-closed, so is its subgroup induced by θ_U. Hence (ii) is a consequence of (i).

If (ii) is satisfied by the pair G, θ then the second part of (iii) is, as a special case of (ii), likewise satisfied. If U is a p-subgroup of G, then the group Γ of automorphisms of U which are induced in U by elements in the normalizer NU of U in G is a subgroup of the group of automorphisms induced in U by automorphisms in θ_U. The latter group of automorphisms of U is strictly p-closed by (ii). Hence Γ is strictly p-closed too. Thus the condition (ii) of Theorem 2.1 is satisfied by G, proving the supersolubility of G. Thus we have shown that (iii) is a consequence of (ii).

Assume next that (iii) is satisfied by the pair G, θ. If K is a θ-admissible normal subgroup of G, then θ induces in G/K a group Σ of automorphisms. Denote by S/K a p-Sylow subgroup of G/K. Then the group of S/K-preserving automorphisms in Σ may be denoted by Σ_S, since it is induced by the automorphisms in θ_S. Denote by P a p-Sylow subgroup of S. Since S/K is a p-Sylow subgroup of G/K, we have $S = KP$ and P is a p-Sylow subgroup of G. Application of (iii) shows that θ_P induces in P a strictly p-closed group

of automorphisms. If σ is an automorphism in θ_S, then $KP = S = S^\sigma = KP^\sigma$, since K is θ-admissible. Hence P^σ is a p-Sylow subgroup of KP. Consequently there exist elements a and b in K and P respectively such that $P^\sigma = P^{ba} = P^a$. Since θ contains every inner automorphism, the automorphism σa^{-1} belongs to θ_P. Since a belongs to K, it induces the identity automorphism in G/K. Thus we have shown that the group of automorphisms induced in S/K by elements in θ_S is an epimorphic image of the group of automorphisms induced by θ_P in P. Since the latter is strictly p-closed, so is the former. Hence Σ_S induces a strictly p-closed group of automorphisms in S/K. Since the Sylow Tower Property of supersoluble groups is inherited by quotient groups, we have shown that (iii) implies the following property:

(iii*) *If K is a θ-admissible normal subgroup of G, then the Sylow Tower Property of supersoluble groups is satisfied by $G/K = H$; and if Σ is the group of automorphisms induced in H by θ, and P is a p-Sylow subgroup of H, then Σ_P induces a strictly p-closed group of automorphisms in P.*

It is not difficult to derive (i) from (iii*). For consider a θ-admissible normal subgroup K of G. If p is the maximal prime divisor of the order of $H = G/K$, then the p-Sylow subgroup P of H is a characteristic p-subgroup of H because of the Sylow Tower Property. Clearly $P \neq 1$; and as a characteristic subgroup of H, P is Σ-admissible, if we denote by Σ the group of automorphisms induced in H by θ. There exists a minimal Σ-admissible subgroup M of H which is part of P. Since θ, and hence Σ, contains every inner automorphism of G and H respectively, M is a normal subgroup of H. Because of (iii*) a strictly p-closed group Γ of automorphisms is induced by Σ in P; and the automorphisms in Γ induce a group Γ^* of automorphisms in the Σ-admissible subgroup M which is strictly p-closed as an epimorphic image of Γ. Since M is a minimal Σ-admissible subgroup of H, Γ^* is irreducible. Since M is a p-group, we may apply Theorem 1.1 to see that M is cyclic of order p. Thus we have established the existence of a cyclic normal Σ-admissible subgroup $M \neq 1$ of H; and this shows that G, θ is a supersoluble pair. This completes the proof of the equivalence of conditions (i) to (iii).

If G, θ is a supersoluble pair, then G and θ are both supersoluble groups, as has been mentioned before, so that both G and θ have the Sylow Tower Property of supersoluble groups. Consider now a subgroup Σ of θ and a Σ-admissible p-subgroup U of G. Then U, Σ is a supersoluble pair too so that their product $U\Sigma$ is a supersoluble group. Thus every maximal subgroup of $U\Sigma$ has index a prime. If V is a maximal Σ-admissible subgroup of U, then $V\Sigma$ is a maximal subgroup of $U\Sigma$. Hence $[U\Sigma : V\Sigma]$ is a prime; and this prime is p since $V < U$. It follows that (iv) is a consequence of (i).

Assume finally the validity of (iv). Since θ contains the group of inner automorphisms of G which is essentially the same as G/ZG, and since θ has the Sylow Tower Property of supersoluble groups, G/ZG has likewise this property. But this implies naturally that G itself enjoys the Sylow Tower Property of supersoluble groups. Consider a p-Sylow subgroup P of G; and

denote by Γ the group of automorphisms induced in P by θ_P. Since ϕP is a characteristic subgroup of P, automorphisms preserving P will also preserve ϕP. Hence ϕP is θ_P-admissible. Denote by Σ a subgroup of θ_P which induces in $P/\phi P$ a group Σ^* of automorphisms whose order is prime to p. Since Σ^* acts on the elementary abelian p-group $P/\phi P$, it is completely reducible (Maschke's Theorem) (**6**, p. 81, Theorem 46). This signifies that every Σ^*-admissible subgroup of $P/\phi P$ possesses in $P/\phi P$ a Σ^*-admissible complement. By (iv), maximal Σ^*-admissible subgroups of $P/\phi P$ have index p in $P/\phi P$. Consequently $P/\phi P$ is the direct product of cyclic Σ^*-admissible subgroups. Since cyclic subgroups of $P/\phi P$ have order p, and since the group of automorphisms of a cyclic group of order p is cyclic of order $p - 1$, it follows that Σ^* is abelian of exponent $p - 1$. We recall the result of P. Hall that an automorphism of the p-group P has order a power of p in case it induces the identity in $P/\phi P$; (§ 1, Lemma (a)). Combining these results we see that a subgroup of Γ whose order is prime to p is abelian of exponent $p - 1$. If the order of Γ is divisible by p, then p is the maximal prime divisor of the order of Γ. By (iv), the group θ, and consequently Γ too, have the Sylow Tower Property of supersoluble groups. This implies in particular the existence of a characteristic p-Sylow subgroup Γ_p of Γ. By Schur's Theorem there exists a complement Δ of Γ_p in Γ. Since $\Delta \simeq \Gamma/\Gamma_p$ is of order prime to p, Δ and consequently Γ/Γ_p is abelian of exponent $p - 1$. Hence Γ is strictly p-closed; and thus we have shown that (iii) is a consequence of (iv). This completes the proof of the equivalence of conditions (i) to (iv).

THEOREM 3.2. *G, θ is a supersoluble pair if, and only if, $G/\phi G, \theta$ is a supersoluble pair.*

Proof. The necessity of our condition is obvious. Assume that $G/\phi G, \theta$ is a supersoluble pair. Since this condition remains valid, if we adjoin the inner automorphisms of G to θ, we may assume without loss in generality that the inner automorphisms of G belong to θ. Consider a θ-admissible normal subgroup K of G. Then $\phi(G/K) = J/K$ where J is the intersection of all those maximal subgroups of G which contain K. This implies in particular that ϕG is part of J and that consequently $K\cdot\phi G/K \leqslant \phi(G/K)$. Thus $(G/K)/\phi(G/K)$ is an epimorphic image of $G/\phi G$. Since $G/\phi G, \theta$ is a supersoluble pair, $(G/K)/\phi(G/K), \theta$ is likewise a supersoluble pair. Let $H = G/K$; and denote by Σ the group of automorphisms induced in G/K by θ. Since $H/\phi H, \Sigma$ is a supersoluble pair, $H/\phi H$ (and the group of automorphisms, induced by Σ in $H/\phi H$) are supersoluble. The supersolubility of $H/\phi H$ implies the supersolubility of H; (**7**, p. 418, Satz 10). If p is the maximal prime divisor of the order of H, then H is p-closed and the p-Sylow subgroup P of H is a characteristic p-subgroup of H. Thus P is in particular Σ-admissible. It is a consequence of Theorem 2.2 (b) that $\phi P = P \cap \phi H$. This implies that Σ induces essentially the same group of automorphisms in $P/\phi P = P/[P \cap \phi H]$ and in $P\cdot\phi H/\phi H$. The latter group is the p-Sylow subgroup of $H/\phi H$ and p is the maximal

prime divisor of the order of $H/\phi H$. Since $H/\phi H$, Σ is a supersoluble pair, Σ induces in $P \cdot \phi H/\phi H$, and hence in $P/\phi P$, a strictly p-closed group of automorphisms. Denote by Γ the group of automorphisms induced in P by automorphisms in Σ; and denote by Γ^* the subgroup of those automorphisms in Γ which induce the identity automorphism in $P/\phi P$. By § 1, Lemma (a), the normal subgroup Γ^* of Γ is a p-group. Since Γ/Γ^* is essentially the same as the group of automorphisms induced in $P/\phi P$ by Σ, and since the latter group is strictly p-closed, we see that Γ is an extension of a p-group by a strictly p-closed group. Hence Γ itself is strictly p-closed. Since $P \neq 1$ is Σ-admissible, there exists a minimal Σ-admissible subgroup M of H which is part of P. Since θ, and hence Σ, contains every inner automorphism, M is a normal subgroup of H. Because of the minimality of M the group Σ induces in M an irreducible group Λ of automorphisms. Since Λ is likewise induced by Γ (because $M \leqslant P$), Λ is strictly p-closed. Since M is a p-group, Theorem 1.1 is applicable. Hence M is cyclic of order p. Thus we have established the existence of a cyclic, θ-admissible, normal subgroup $M \neq 1$ of G/K. Hence G, θ is a supersoluble pair, q.e.d.

4. The results obtained in § 3 will now be applied to the problem of supersoluble immersion.

THEOREM 4.1. *The following properties of the normal subgroup K of G are equivalent:*

(i) *K is supersolubly immersed in G.*
(ii) *$K/\phi K$ is supersolubly immersed in $G/\phi K$.*
(iii) *If U is a p-subgroup of K, then NU/CU is strictly p-closed.*
(iv) *K has the Sylow Tower Property of supersoluble groups; and if P is a p-Sylow subgroup of K, then NP/CP is strictly p-closed.*
(v) *G/CK has the Sylow Tower Property of supersoluble groups; and if P is a p-Sylow subgroup of K, S a subgroup of NP, then maximal S-normalized subgroups of P have index p in P.*

Proof. Denote by θ the group of automorphisms induced in K by elements in G. Then θ is essentially the same as G/CK. If U is a subgroup of K, then θ_U is just the group of all automorphisms of U which are induced in U by elements in the normalizer NU of U in G; and this shows that θ_U and NU/CU are essentially the same. Note finally that K is supersolubly immersed in G if, and only if, K, θ is a supersoluble pair. Since the inner automorphisms of K are clearly contained in θ, Theorems 3.1 and 3.2 may be applied, and a fairly obvious translation of these results proves the equivalence of properties (i) to (v).

LEMMA 4.2. *If K is a normal subgroup of G, and if the subgroup S of G is minimal with respect to the property $G = KS$, then $K \cap S \leqslant \phi S$, and $S/\phi S$ is an epimorphic image of G/K. In particular S is supersoluble in case G/K is supersoluble.*

Proof. Consider a maximal subgroup T of S. If $K \cap S$ were not part of T, then we could deduce from the maximality of T (and the normality of $K \cap S$ in S) that $S = (K \cap S)T$. Consequently

$$G = KS = K(K \cap S)T = KT.$$

But $T < S$, contradicting the minimality of S. Thus we see that $K \cap S$ is part of every maximal subgroup of S; in other words: $K \cap S \leq \phi S$. Next we note the isomorphy $G/K \simeq S/(K \cap S)$. From $K \cap S \leq \phi S$ we may deduce therefore that $S/\phi S$ is an epimorphic image of G/K. Thus supersolubility of G/K implies the supersolubility of $S/\phi S$; and the latter implies, by Huppert's Theorem, the supersolubility of S.

Remark. This lemma is, naturally, well known. It has been appended for the convenience of the reader. Note that the first part of the lemma has many applications of the type given in its second part, since there exist many group theoretical properties which, when satisfied by a group, are satisfied by its epimorphic images, and which, when satisfied modulo the Frattini subgroup, are satisfied by the group itself; for instance, nilpotency, dispersion, etc.

THEOREM 4.3. *The normal subgroup K of G is supersolubly immersed in G if, and only if,*

(a) *G induces in K a supersoluble group of automorphisms and*

(b) *the supersolubility of the subgroup S of G implies the supersolubility of KS.*

Proof. The necessity of these conditions we have pointed out before. If the conditions (a) and (b) are satisfied by the normal subgroup K of G, then we select among the subgroups X of G satisfying $G = X \cdot CK$ a minimal one, say S. Since the group of automorphisms, induced in K by elements in G, is essentially the same as G/CK, this group is supersoluble by (a). Application of Lemma 4.2 shows the supersolubility of S. Application of (b) shows the supersolubility of KS. Denote now by θ the group of automorphisms induced in K by elements in G. Because of $G = S \cdot CK$ the elements in KS induce in K the same group θ of automorphisms. Since KS is supersoluble, the pair K, θ is a supersoluble pair. But then clearly K is supersolubly immersed in G, q.e.d.

We have pointed out before that *the product $\Sigma_i G$ of all the supersolubly immersed normal subgroups of G* is itself a supersolubly immersed characteristic subgroup of G. If we denote by $\Sigma_0 G$ *the product of all normal subgroups X of G such that XS is supersoluble whenever S is a supersoluble subgroup of G*, then $\Sigma_0 G$ is a characteristic subgroup of G satisfying the same property (b) of Theorem 4.3. Denote finally by ΣG *the intersection of all normal subgroups X of G with supersoluble G/X*. Since direct products and subgroups of supersoluble groups are themselves supersoluble, ΣG is a characteristic subgroup of G with supersoluble quotient group $G/\Sigma G$.

Corollary 4.4. $\Sigma_i G = \Sigma_0 G \cap C\Sigma G$.

Proof. If K is a supersolubly immersed normal subgroup of G, then we deduce $K \leqslant \Sigma_0 G$ from Theorem 4.3 (b) and $\Sigma G \leqslant CK$ from Theorem 4.3 (a). The latter inequality implies $K \leqslant C\Sigma G$. Thus we have shown that

$$\Sigma_i G \leqslant \Sigma_0 G \cap C\Sigma G.$$

Let $D = \Sigma_0 G \cap C\Sigma G$. Then D is a characteristic subgroup of G which satisfies $D \leqslant C\Sigma G$ and hence $\Sigma G \leqslant CD$, implying the validity of condition (a) of Theorem 4.3. From $D \leqslant \Sigma_0 G$ we deduce the validity of condition (b) of Theorem 4.3. It follows that D is supersolubly immersed in G. Hence $D \leqslant \Sigma_i G$, completing the proof.

It is worth noting in this context that, in general, $\Sigma_i G < \Sigma_0 G$, and that products of supersoluble normal subgroups will, in general, not be supersoluble.

Slightly generalizing the concept of a supersoluble pair we term the pair G, θ (for θ a group of automorphisms of the group G) *an almost supersoluble pair, if G, Σ is, for every supersoluble subgroup Σ of θ, a supersoluble pair.* If the pair G, θ is almost supersoluble, then G, Σ is a supersoluble pair for every Sylow subgroup Σ of θ. The converse is false, as may be seen from the following

Example. Let p be a prime, q an odd prime divisor of $p - 1$ (for instance, $p = 7$, $q = 3$). Then $2q$ is a factor of $p - 1$. There exists one and essentially only one non-abelian group θ of order $2q$; and θ possesses a normal subgroup of order q and index 2, its only proper normal subgroup. It follows among other things that θ is supersoluble. There exists an elementary abelian p-group A of order p^{2q}; and there exists a group of automorphisms of A which is isomorphic to θ and which we shall denote by θ. Since Sylow subgroups of θ are cyclic of order q or 2, they are strictly p-closed. Hence every pair A, Σ, for Σ a Sylow subgroup of θ, is supersoluble. But θ itself is not strictly p-closed, though it is a group of automorphisms of the p-group A. Hence A, θ is not a supersoluble pair.

The connection between the concept of an almost supersoluble pair and our preceding discussion is effected by the following fairly obvious remark: If K is a normal subgroup of G and θ the group of automorphisms induced in K by the elements in G, then the pair K, θ is almost supersoluble if, and only if, KS is supersoluble whenever S is a supersoluble subgroup of G (Theorem 4.3).

References

1. R. Baer, *Supersoluble groups*, Proc. Amer. Math. Soc., *6* (1955), 16–32.
2. ———, *Classes of finite groups and their properties*, Illinois Jour. Math., *1* (1957), 115–187.
3. ———, *Closure and dispersion*, Illinois Jour. Math., *3* (1959), 000–000.
4. P. Hall, *A contribution to the theory of groups of prime power order*, Proc. Lond. Math. Soc., *36* (1932), 29–95.
5. B. Huppert, *Normalteiler und maximale Untergruppen endlicher Gruppen*, Math. Zeit., *60* (1954), 409–34.
6. N. Jacobson, *The Theory of Rings* (New York, 1943).
7. H. Zassenhaus, *Lehrbuch der Gruppentheorie, I* (Leipzig und Berlin, 1937).

Frankfurt am Main

NON-LINEAR RECURSIVE SEQUENCES

ELBERT A. WALKER

The purpose of this paper is to investigate non-linear recursive sequences of maximum length with elements from $GF(2)$. In particular, the question of whether or not a recursive sequence of maximum length can be equal to its dual is settled. This question, as far as the author knows, was originally asked by Rosser. Part I contains the necessary background for Part II, and in the main is a condensation of some unpublished work (1955) of W. A. Blankinship and R. P. Dilworth.

PART I

1. Let $GF(2)$ be the field with two elements, let n be a positive integer, let $\mathfrak{S}$ be the Cartesian product of n copies of $GF(2)$, and let f be a mapping from $\mathfrak{S}$ into $GF(2)$. A sequence $a_1, a_2, a_3, \ldots$ of elements in $GF(2)$ is said to be ***recursively generated*** by f if

$$a_{n+i} = f(a_i, a_{i+1}, \ldots, a_{n+i-1}) \qquad \text{for i} = 1, 2, 3, \ldots$$

f is called a ***recursion*** or a ***rule of generation***. The sequence $a_1, a_2, a_3, \ldots$ is called a ***recursive sequence*** of span $\leq n$. It is of span n if in addition it is not of span $\leq n-1$. The elements of $\mathfrak{S}$ will be called *patterns*, or *n-bit words*, and the elements of $GF(2)$ will sometimes be called *bits*, and denoted by 0, 1. The mapping f induces a mapping F of $\mathfrak{S}$ into $\mathfrak{S}$ in the following way. If $S = (a_1, a_2, \ldots, a_n)$ is in $\mathfrak{S}$, let $F(S) = (a_2, a_3, \ldots, a_n, f(a_1, \ldots, a_n))$. Hence with the mapping f of $\mathfrak{S}$ into $GF(2)$, we associate the mapping F of $\mathfrak{S}$ into $\mathfrak{S}$, and F uniquely determines f. Distinct mappings f_1 and f_2 of $\mathfrak{S}$ into $GF(2)$ induce distinct mappings F_1 and F_2 of $\mathfrak{S}$ into $\mathfrak{S}$. If F is one-to-one, then f is said to be a ***non-singular*** recursion. Otherwise f is called ***singular***. All recursions considered here will be assumed to be non-singular unless otherwise stated.

2. Let f be a recursion, let F be the mapping of $\mathfrak{S}$ into $\mathfrak{S}$ determined by f and let S_1 be in $\mathfrak{S}$. Let $S_i = F(S_{i-1})$, $i = 2, 3, 4, \ldots$. Since $\mathfrak{S}$ is finite, there exists a smallest positive integer m such that $S_{m+1} = S_1$. Thus f generates a cycle of elements in $\mathfrak{S}$, namely $(S_1, S_2, \ldots, S_m)$. If T_1 is some element in $\mathfrak{S}$ not in the cycle $(S_1, S_2, \ldots S_m)$, f generates another cycle $(T_1, T_2, \ldots T_{m'})$ and these two cycles are disjoint. Continuing in this manner, $\mathfrak{S}$ is decomposed into disjoint cycles by f. Of course the cycle $(S_1, S_2, \ldots, S_m)$ is con-

Received October 17, 1958.

sidered the same as $(S_2, S_3, \dots, S_m, S_1)$. This collection of cycles determined by f is denoted by C, and it is easy to see that f is uniquely determined by C. The system C of cycles is called the cyclic structure of f. Since F is one-to-one, it is onto, and so is a permutation of $\mathfrak{S}$. If this permutation is decomposed into the product of disjoint cycles, this collection of cycles is identical with C. Thus the cyclic structure of the permutation F is identical with the cyclic structure of f. The sum of the lengths of the cycles in C is 2^n, where n is the span of f. If C consists of just one cycle, that cycle is said to be a *maximal* cycle.

3. Any mapping f of $\mathfrak{S}$ into $GF(2)$, singular or non-singular, can be represented uniquely as a polynomial in $x_1, x_2, \dots, x_n$ with coefficients in $GF(2)$. If f is linear in x_1,

$$f(a_1, a_2, \dots, a_n) = f(a_1 + 1, a_2, \dots, a_n) + 1.$$

Hence f linear in x_1 implies f is non-singular. Assume f is non-singular. Then $f(1, 0, 0, 0, \dots, 0) = 1 + f(0, 0, \dots, 0)$ so that the term x_1 appears in the polynomial representing f. $f(1, 1, 1, \dots, 1) = 1 + f(0, 1, 1, \dots, 1)$ so that if the polynomial representing f has any non-linear terms with x_1 as a factor, it has an even number of them. If it has at least two, let

$$x_1x_{i_1} \dots x_{i_r} \quad \text{and} \quad x_1x_{j_1}x_{j_2} \dots x_{j_s}$$

be distinct and let the first have smallest possible degree. The sets $\{i_1, i_2, \dots, i_r\}$ and $\{j_1, j_2, \dots, j_s\}$ are distinct. There is a j_k not in $\{i_1, i_2, \dots, i_r\}$. Let $S = (a_1, a_2, \dots, a_n)$ be the element of $\mathfrak{S}$ whose first co-ordinate is 1, whose $i_1, i_2, \dots, i_r$ co-ordinates are 1, and the rest of whose co-orinates are 0. Let $S' = (a_1 + 1, a_2, a_3, \dots, a_n)$. Then $F(S) = F(S')$, and F is singular. Hence f is linear in x_1. Thus f is non-singluar if, and only if, f is dependent on x_1 and linear in x_1, and $f(x_1, x_2, \dots, x_n) = x_1 + f_1(x_2, x_3, \dots, x_n)$, where f_1 is a polynomial in $x_2, x_3, \dots, x_n$.

4. Let $S = (a_1, a_2, \dots, a_n)$ be a pattern in $\mathfrak{S}$, and let $\bar{S} = (a_1 + 1, a_2, \dots, a_n)$. Let f_S be the mapping from $\mathfrak{S}$ into $GF(2)$ that is 1 only at S and $\bar{S}$. Explicitly,

$$f_S(x_1, \dots, x_n) = \prod_{i=2}^{n} (1 + a_i + x_i).$$

Note that $f_S = f_{\bar{S}}$. Let C be the system of cycles of f. Suppose S and $\bar{S}$ are on distinct cycles in C. Let $(S_1, S_2, \dots, S_k)$ be the cycle containing S, and let $(T_1, T_2, \dots, T_m)$ be the cycle containing $\bar{S}$. For convenience, let $S_1 = S$ and $T_1 = \bar{S}$. Then the system C' of cycles of $f + f_S$ consists of the cycle $(S_1, T_2, T_3, \dots, T_m, T_1, S_2, S_3, \dots, S_k)$ and the remaining cycles of C unchanged. Suppose S and $\bar{S}$ are on the same cycle $(S_1, S_2, \dots, S_k)$ in C. Let $S_1 = S$ and $S_r = \bar{S}$. Then the system C' of cycles of $f + f_S$ consists of the cycles $(S_1, S_{r+1}, \dots, S_k)$, $(S_r, S_2, \dots, S_{r-1})$, and the remaining cycles of C unchanged.

5. Let f be a recursion that generates a cycle $C_1 = (S_1, S_2, \ldots, S_k)$. Suppose this cycle C_1 has the property that if it contains the pattern $(a_1, a_2, \ldots, a_n)$ then it contains $(a_1 + 1, a_2, a_3, \ldots, a_n)$. Let $(b_1, b_2, \ldots, b_n)$ be any pattern. C_1 contains a pattern ending in b_1, $(a_1, a_2, \ldots, a_{n-1}, b_1)$. If this pattern is not followed in C_1 by $(a_2, a_3, \ldots, a_{n-1}, b_1, b_2)$, then the pattern $(a_1 + 1, a_2, \ldots, a_{n-1}, b_1)$, which is in C_1, is followed by $(a_2, a_3, \ldots, a_{n-1}, b_1, b_2)$. Continuing in this manner, one gets the pattern $(b_1, b_2, \ldots, b_n)$ in C_1. Hence the system of cycles of f consists simply of the one cycle C_1. If a recursion f generates more than one cycle, then every cycle it generates has the property that it contains a pattern S such that it does not contain $\bar{S}$. Therefore $f + f_S$ generates one less cycle than does f. In general, if f generates k cycles, then there exist $k - 1$ patterns $S_1, \ldots, S_{k-1}$ such that

$$f + \sum_{i=1}^{k-1} f_{S_i}$$

generates just one cycle.

Part II

An unsolved problem concerning non-linear recursive sequences is that of finding a large class of recursions which generate maximal cycles. It is known **(1)** that the number of such recursions of span n is

$$2^{2^{n-1}-n}.$$

We begin this section by deriving some elementary properties a recursion must have if it generates a maximal cycle. Later we define and investigate the reverse, the dual ,and the reverse-dual of a recursion.

1. Let f be a recursion of span n that generates a maximal cycle. Then

$$f(x_1, x_2, \ldots, x_n) = 1 + x_1 + \phi(x_2, x_3, \ldots, x_n),$$

where $\phi\ (x_2, x_3, \ldots, x_n)$ is a polynomial with no constant term.

Proof. Since f generates a maximal cycle, every n-bit word must occur in that cycle. Thus the n-bit word $(0, 0, \ldots, 0)$ is not a rut, that is $f(0, 0, \ldots, 0) = 1$.

2. Let f be a recursion of span $n > 1$ that generates a maximal cycle. Then the polynomial $f(x_1, \ldots, x_n)$ that represents f does not contain all the linear terms $x_2, \ldots, x_n$.

Proof. The n-bit word $(0, 0, \ldots, 0)$ is followed by a 1. If $f(x_1, \ldots, x_n)$ has the term x_n, then the $n + 1$-bit word $(0, 0, \ldots, 0, 1)$ is followed by 0, and if it also has the term x_{n-1}, this pattern is followed by 0, etc. If $f(x_1, \ldots, x_n)$ has all the linear terms $x_2, \ldots, x_n$, we get the sequence $0, 0, \ldots, 0, 1, 0, 0, \ldots, 0$ since $f(x_1, \ldots, x_n)$ contains the linear term x_1. But if $n > 1$, this

implies that f does not generate a maximal cycle. Therefore $f(x_1, \ldots, x_n)$ does not contain all those terms.

3. Let f be a recursion of span n that generates a maximal cycle. Then the polynomial that represents f has an even number of terms.

Proof. The n-bit word $(1, 1, \ldots, 1)$ is followed by a 0. Hence $f(1, 1, \ldots, 1) = 0$ and so $f(x_1, \ldots, x_n)$ has an even number of terms.

4. Definitions. One cycle is the *reverse* of another if either cycle can be obtained from the other by taking the bits in reserve order. One cycle is the *dual* of another if either can be found from the other by replacing all 0's by 1's and all 1's by 0's. One cycle is the *reverse-dual* of another if it is the reverse of the dual (the same as the dual of the reverse) of the other. The recursion corresponding to the reverses of the cycles generated by f is called the *reverse* of f, and is denoted by Rf. The recursion corresponding to the duals of the cycles generated by f is called the *dual* of f, and is denoted by Df. The recursion corresponding to the reverse-dual of the cycles of f is called the *reverse-dual* of f, and is denoted by RDf.

5. It is fairly clear that the cyclic structure of f, Rf, Df, and RDf are the same as far as the number of cycles in each and the lengths of cycles in each are concerned. In particular, if f generates a maximal cycle, then so do Df, Rf, and RDf.

6. Since every recursion f can be represented by a polynomial, it is of some interest to determine the polynomials representing Rf, Df, and RDf in terms of the one representing f. A moment's reflection shows that if

$$f(x_1, x_2, \ldots, x_n) = x_1 + f_1(x_2, \ldots, x_n)$$

then

$$Rf(x_1, \ldots, x_n) = x_1 + f_1(x_n, \ldots, x_2),$$

and

$$Df(x_1, \ldots, x_n) = x_1 + f_1(1 + x_2, 1 + x_3, \ldots, 1 + x_n).$$

From these follow then that

$$RDf(x_1, \ldots, x_n) = x_1 + f_1(x_n + 1, \ldots, x_2 + 1).$$

7. Since $f(x_1, \ldots, x_n) = x_1 + f_1(x_2, \ldots, x_n)$ implies that

$$Rf(x_1, \ldots, x_n) = x_1 + f_1(x_n, \ldots, x_2),$$

we see that f and Rf agree on those patterns which are symmetric in the last $(n - 1)$ bits.

8. Suppose a cycle $(S_0, S_1, \ldots, S_{k-1})$ is the same as its reverse. Suppose

this cycle contains a pattern which is its own reverse, and for convenience let it be S_0. Let S_i' be the reverse of the pattern S_i. Then $S_0 = S_0'$, $S_1 = S_{k-1}'$, $\ldots, S_j = S_{k-j}', \ldots$. There is at most one j such that $j = k - j$, namely $j = \frac{1}{2}k$. Therefore, if a cycle is the same as its reverse, then it contains at most two patterns which are their own reverses. If k is odd, there is at most one such pattern, and as a matter of fact, exactly one such pattern. It is possible for a cycle of even length to be its own reverse and contain no pattern which is its own reverse. For example, the cycle ((01), (10)) is such a cycle. Now, using these facts we can prove the following theorem.

THEOREM. *If $f = Rf$, then f generates at least $2^{[\frac{1}{2}n+\frac{1}{2}]-1}$ cycles. If $n > 3$, then f does not generate a maximal cycle.*

Proof. There are $2^{[\frac{1}{2}n+\frac{1}{2}]}$ n-bit words which are their own reverses. Since $f = Rf$, a cycle which contains one of these n-bit words is its own reverse. But a cycle that is its own reverse can contain at most two n-bit words that are their own reverses. Hence there must be at least $2^{[\frac{1}{2}n+\frac{1}{2}]-1}$ cycles generated by f. If $n > 3$, $2^{[\frac{1}{2}n+\frac{1}{2}]-1} > 2$, so that f does not generate a maximal cycle.

9. We are now going to settle the question as to whether or not a cycle of maximal length can be equal to its dual. It has just been shown that if $n > 3$, a maximal cycle of span n is not equal to its reverse. The corresponding statement is true for the dual of a maximal cycle, but the proof of it is a little more complicated. We begin with a lemma, of which no proof seems readily available in the literature.

LEMMA. *If $f(x_1, \ldots, x_n) = x_1$, then the number of cycles generated by f is even, for $n > 2$.*

Proof. If $a_1a_2 \ldots a_n$ is any pattern, then the sequence obtained beginning with this pattern is $a_1a_2 \ldots a_na_1a_2 \ldots a_n \ldots$. Therefore to compute the number of cycles generated by f is the same problem as computing the number of strings of beads of length n that can be constructed using two kinds of beads, where two strings are considered the same if one is a rotation of the other. It is easily verified that this number is

$$G(n) = \sum_{d \mid n} \frac{F(d)}{d},$$

where $F(1) = 2$ and

$$F(k) = 2^k - \sum_{\substack{r \mid k \\ r<k}} F(r).$$

$F(d)$ is, in fact, the number of patterns that are equal to themselves at slides of multiples of d only. Hence such a pattern and its slides contain exactly d distinct patterns, from which it follows that $d^{-1}F(d)$ is the number of strings of beads n long of this nature. Summing over all divisors d of n yields the

total number of strings of beads. This number $G(n)$ we wish to show is even. Observing that

$$\sum_{d|n} F(d) = 2^n$$

and applying the Möbius inversion formula yields

$$F(n) = \sum_{d|n} \mu(n/d)2^d.$$

Therefore we get

$$G(n) = \sum_{d|n} \frac{F(d)}{d} = \sum_{d|n} \sum_{r|d} \frac{\mu(d/r)2^r}{d}.$$

If n is odd we see immediately that $F(d)$ and hence $d^{-1}F(d)$ is even for all $d|n$. Thus we need consider only the case where

$$d = 2^a p_1^{a_1} p_2^{a_2} \ldots p_s^{a_s}$$

where $a > 0$, and $p_1, p_2, \ldots, p_s$ are distinct odd primes.* Now

$$F(d) = \sum_{r|d} \mu(r)2^{d/r}$$

and each non-vanishing term in the sum has a factor

$$2^{2^{a-1}p_1^{a_1-1}p_2^{a_2-1}\ldots p_s^{a_s-1}}.$$

Hence

$$F(d) = m2^{2^{a-1}p_1^{a_1-1}\ldots p_s^{a_s-1}}$$

where m is an integer. Since $d|F(d)$,

$$p_1^{a_1}p_2^{a_2}\ldots p_s^{a_s}|m.$$

Put

$$m = up_1^{a_1}p_2^{a_2}\ldots p_s^{a_s},$$

where u is an integer. Then

$$\frac{F(d)}{d} = u2^{(2^{a-1}p_1^{a_1-1}p_2^{a_2-1}\ldots p_s^{a_s-1}-a)}.$$

A necessary condition that $d^{-1}F(d)$ be odd is $a_1 = a_2 = \ldots = a_s = 1$ and $a = 1$ or 2; that is, if $d = 2p_1p_2\ldots p_s$ or $d = 4p_1p_2\ldots p_s$. We show in each of these cases that $d^{-1}F(d)$ is odd. Let $d = 2p_1p_2\ldots p_s$. Then

$$F(d) = 2^d \pm 2^{d_1} \pm 2^{d_2} \pm \ldots \pm 2$$

where $d, d_1, d_2, \ldots,$ are all the divisors of d. Hence $F(d)$ is twice an odd number and since $d|F(d)$, $d^{-1}F(d)$ is odd. If $d = 4p_1p_2\ldots p_s$ then since

$$F(d) = \sum_{r|d} \mu(r)2^{d/r}$$

---

*The author wishes to thank the referee for furnishing a correct proof for this case.

and $\mu(r) = 0$ if 4 divides r,

$$F(d) = 2^d \pm 2^{d_1} \pm 2^{d_2} \pm \ldots \pm 2^2$$

where $d, d_1, d_2, \ldots, 2$, are all the *even* divisors of d. Hence $F(d)$ is four times an odd number and since $d|F(d)$, $d^{-1}F(d)$ is odd. Now let

$$n = 2^a p_1^{a_1} p_2^{a_2} \ldots p_s^{a_s}.$$

Since $n > 2$, either $s > 0$ or $a > 1$. If $s = 0$, $n = 2^a$, $a > 1$ and

$$G(n) = \sum_{d|n} \frac{F(d)}{d} = \frac{F(1)}{1} + \frac{F(2)}{2} + \frac{F(4)}{4} + \sum_{r=3}^{a} \frac{F(2^r)}{2^r}$$
$$= 2 + 1 + 3 + \sum \text{ (even numbers)}.$$

Hence $G(n)$ is even. If $s > 0$, $a = 0$, then each term of

$$\sum_{d|n} \frac{F(d)}{d}$$

is even. If $s > 0$, $a = 1$, then the divisors d, for which $d^{-1}F(d)$ is odd, are the numbers $d = 2q_1q_2 \ldots q_r$, where $q_1, q_2, \ldots, q_r$ is a subset of $p_1, p_2, \ldots, p_s$. The number of such divisors is 2^s so that $G(n)$ is even. Finally, if $s > 0, a > 1$, the divisors d, for which $d^{-1}F(d)$ is odd, are of the forms $d = 2q_1q_2 \ldots q_r$ or $d = 4q_1q_2 \ldots q_r$ where $q_1, q_2, \ldots, q_r$ is a subset of $p_1, p_2, \ldots, p_s$. The number of such divisors is 2^{s+1}. Hence, again $G(n)$ is even.

10. Let f and g be recursions of span n. If f and g disagree on the n-bit word S, then from 4, Part I, we see that $f + f_S$ and g agree on S and $\bar{S}$ and on all patterns on which f and g agree. Thus the recursion f may be changed into the recursion g by adding a suitable set of f_S's to f. From 5, Part I, we see that adding h_S to any recursion h changes the parity of the number of cycles generated by h. If $S = (a_1, a_2, \ldots, a_n)$ then

$$h_S(x_1, \ldots, x_n) = \prod_{i=2}^{n} (1 + a_i + x_i),$$

and adding h_S to h adds the term $x_2x_3 \ldots x_n$, among other terms, to the polynomial representing h. If one adds an odd number of h_S's to h one adds the term $x_2x_3 \ldots x_n$, among other terms, to the polynomial representing h. Now let f be the recursion such that $f(x_1, \ldots, x_n) = x_1$ and let g be any recursion of the same span that generates an odd number of cycles. If $n > 2$, f generates an even number of cycles, so to change f into g requires the adding of an odd number of f_S's to f, and hence the adding of the term $x_2x_3 \ldots x_n$, among other terms to the polynomial representing f, which is x_1. Hence the polynomial representing g has the term $x_2x_3 \ldots x_n$. Conversely, if g is any recursion of span n such that the polynomial representing it has the term $x_2x_3 \ldots x_n$, then one must add an odd number of f_S's to f to get g, and this implies that g generates an odd number of cycles. These remarks we sum up in the following theorem.

THEOREM. *A recursion of span $n > 2$ generates an odd number of cycles if, and only if, the polynomial representing it has the term $x_2x_3 \ldots x_n$.*

COROLLARY. *If a recursion of span $n > 2$ generates a maximal cycle, then the polynomial representing it has the term $x_2x_3 \ldots x_n$.*

11. We are now in a position to prove that if $n > 2$ and f generates a maximal cycle, then $f \neq Df$. In fact, we will prove a more general result.

THEOREM. *If $n > 2$ and a recursion f generates an odd number of cycles, then $f \neq Df$.*

Proof. From 6, Part II, it is easy to see that if the polynomial representing f contains a term

$$x_{i_1}x_{i_2} \ldots x_{i_r},$$

then this term is a term of the polynomial representing Df if, and only if,

$$x_{i_1}x_{i_2} \ldots x_{i_r}$$

is a factor of an odd number of terms of the polynomial representing f. Since f generates an odd number of cycles, the polynomial representing it contains the term $x_2x_3 \ldots x_n$. If that polynomial contains a term besides 1, x_1, and $x_2x_3 \ldots x_n$, then it contains a term which is a factor of only itself and $x_2x_3 \ldots x_n$. That term then is a factor of an even number of terms, and therefore is not a term of the polynomial representing Df. If

$$f(x_1, \ldots, x_n) = 1 + x_1 + x_2x_3 \ldots x_n \text{ or } x_1 + x_2x_3 \ldots x_n$$

then

$$Df(x_1, \ldots, x_n) = 1 + x_1 + (1 + x_2)(1 + x_3) \ldots (1 + x_n)$$
$$\text{or } x_1 + (1 + x_2)(1 + x_3) \ldots (1 + x_n),$$

and is obviously not the same as $f(x_1, \ldots, x_n)$. Hence in any case, $f \neq Df$.

COROLLARY. *If $n > 2$ and f is a recursion generating a maximal cycle, then $f \neq Df$.*

12. THEOREM. *If n is even and if f is a recursion of span $n > 2$ which generates an odd number of cycles, then $Rf \neq Df$, and hence $f \neq RDf$.*

Proof. From 6, Part II, it follows that the polynomials representing the recursions f and Rf have the same structure with regard to the number of terms which are the product of a given number of variables. There are $n - 1$ possible terms which are the product of $n - 2$ variables, namely $x_3x_4 \ldots x_n$, $x_2x_4 \ldots x_n, \ldots, x_2x_3 \ldots x_{n-1}$. Since the polynomial representing f contains the term $x_2x_3 \ldots x_n$, we see from the proof of the theorem in 11, Part II, that the polynomial representing Df contains precisely those terms which are the product of $n - 2$ variables that the polynomial representing f does

not contain. Thus for Rf to be equal to Df it is necessary that $n - 1$ be even. If n is even then $n - 1$ is odd so that $Rf \neq Df$ and $f \neq RDf$.

COROLLARY. *If $n > 2$ is even and f generates a maximal cycle then $Rf \neq Df$ and $RDf \neq f$.*

For n odd it can happen that $Rf = Df$. It happens in the case $n = 5$, as shown by the polynomial

$$f(x_1, x_2, x_3, x_4, x_5) = 1 + x_1 + x_4 + x_5 + x_4x_5 + x_2x_3x_5 + x_2x_3x_4 + x_2x_3x_4x_5.$$

REFERENCE

1. N. de Bruijn, *A combinatorial problem*, Koninklijke Nederlandse Akademie van Wetenschappen, Proceedings, *49* (Part 2) (1946), 758–64.

New Mexico State University

EXTREMAL PROPERTIES OF HERMITIAN MATRICES. II

M. MARCUS, B. N. MOYLS, AND R. WESTWICK

1. Introduction. Let H be an n-square Hermitian matrix with eigenvalues $h_1 \geqslant h_2 \geqslant \ldots \geqslant h_n$. Fan **(2)** showed that

$$(1) \qquad \begin{cases} \max \sum_{j=1}^{k} (Hx_j, x_j) = \sum_{j=1}^{k} h_j, \\ \min \sum_{j=1}^{k} (Hx_j, x_j) = \sum_{j=1}^{k} h_{n-k+j} \end{cases}$$

$k = 1, 2, \ldots, n$, where the max and min are taken over all sets of k orthonormal (o.n.) vectors in unitary n-space V_n. Marcus and McGregor **(3)** have generalized this result in the case that H is non-negative Hermitian. For vectors $x_1, \ldots, x_r$, $r \leqslant n$, in V_n, let $x_1 \wedge x_2 \wedge \ldots \wedge x_r$ denote the Grassmann exterior product of the x_i; it is a vector in V_m, where

$$m = \binom{n}{r}.$$

The rth compound of H is a Hermitian transformation of V_m defined by

$$C_r(H)\, x_1 \wedge \ldots \wedge x_r = Hx_1 \wedge \ldots \wedge Hx_r.$$

For $1 \leqslant r \leqslant k \leqslant n$, denote by Q_{kr} the set of $\binom{k}{r}$ distinct sequences $w = \{i_1, \ldots, i_r\}$ of integers such that $1 \leqslant i_1 < \ldots < i_r \leqslant k$. For a set of vectors $x_1, \ldots, x_k$ in V_n, set

$$x_w = x_{i_1} \wedge \ldots \wedge x_{i_r}.$$

Let

$$(2) \qquad g = g(x_1, \ldots, x_k) = \sum_{w \in Q_{kr}} (C_r(H)x_w, x_w),$$

and let $E_r(a_1, \ldots, a_k)$ be the rth elementary symmetric function of the numbers $a_1, \ldots, a_k$. Marcus and McGregor showed that

$$(3) \qquad \begin{cases} \max g = E_r(h_1, \ldots, h_k) \\ \min g = E_r(h_{n-k+1}, \ldots, h_n), \end{cases}$$

where the max and min are taken over all sets of k o.n. vectors $x_1, \ldots, x_k$ in V_n. This result reduces to (1) when $r = 1$. In the present note we extend this result to the case where H is an arbitrary Hermitian matrix.

Received July 23, 1958. The work of the first author was supported in part by United States National Science Foundation Research Grant NSF-G 5416; that of the second author by the United States Air Force Office of Scientific Research, Air Research and Development Command; that of the third author by the National Research Council of Canada.

2. Results.

THEOREM. *Let $1 \leqslant r \leqslant k \leqslant n$ and let H be a Hermitian matrix with eigenvalues $h_1 \geqslant \ldots \geqslant h_n$. Then*

$$(4)\qquad \begin{cases} \max g = \max\limits_{0 \leqslant s \leqslant k} E_r(h_1, \ldots, h_s, h_{n-k+s+1}, \ldots, h_n)^* \\ \min g = \min\limits_{0 \leqslant s \leqslant k} E_r(h_1, \ldots, h_s, h_{n-k+s+1}, \ldots, h_n), \end{cases}$$

where the max and min of g are taken over all sets of k o.n. *vectors $x_1, \ldots, x_k$ in V_n.*

Proof. Let $L = L(x_1, \ldots, x_k)$ denote the subspace spanned by the o.n. vectors $x_1, \ldots, x_k$; and let P be the orthogonal projection of V_n into L. Then, since P is Hermitian,

$$\begin{aligned} g(x_1, \ldots, x_k) &= \sum_{w \in Q_{kr}} (C_r(H)x_w, C_r(P)x_w) \\ &= \sum_{w \in Q_{kr}} (C_r(PH)x_w, x_w) \\ &= \text{trace of } C_r(A) \\ &= E_r(\lambda_1, \ldots, \lambda_k), \end{aligned}$$

where A is the Hermitian transformation PH restricted to L, and $\lambda_1 \geqslant \ldots \geqslant \lambda_k$ are the eigenvalues of A. It is known (1, p. 33) that for $1 \leqslant j \leqslant k$,

$$(5)\qquad h_j \geqslant \lambda_j \geqslant h_{n-k+j}.$$

Let $R_k(h)$ be the set of real k-tuples $\lambda = (\lambda_1, \ldots, \lambda_k)$, $\lambda_1 \geqslant \ldots \geqslant \lambda_k$, satisfying the inequalities (5). Thus the values of g are bounded by the extreme values of $E_r(\lambda) = E_r(\lambda_1, \ldots, \lambda_k)$ as λ ranges over $R_k(h)$. We shall discuss the maximum value of $E_r(\lambda)$ in the following lemmas. Corresponding results hold for the minimum. For the moment we restrict ourselves to the case in which the h_j are distinct.

LEMMA 1. *Let $h_1 > \ldots > h_n$ be given real numbers. Let $1 \leqslant r \leqslant k \leqslant n$, and let*

$$(6)\qquad \gamma = \max_{\lambda \in R_k(h)} E_r(\lambda).$$

Then there exists $\mu \in R_k(h)$ such that

$$(7)\qquad E_r(\mu) = \gamma$$

and $\mu_1 > \ldots > \mu_k$.

Proof. When $r = 1$, the unique solution of (7) is: $\mu_j = h_j$, $j = 1, \ldots, k$. Hence suppose that $2 \leqslant r \leqslant k$.

Let $T_{kj}(h)$ be the set of $\lambda = (\lambda_1, \ldots, \lambda_k) \in R_k(h)$ such that $E_r(\lambda) = \gamma$ and $\lambda_1 > \ldots > \lambda_j$. Then $T_{k1}(h)$ is not void by the continuity of the elemen-

*If $s = 0$ (or k) the initial (or terminal) segment is missing.

tary symmetric functions. Let m be the least integer such that $T_{km}(h)$ is not void. Then m must equal k for, if not, we shall show that there exists $\nu \in T_{k,m+1}(h)$. Suppose then that $\mu \in T_{km}(h)$, where

$$\mu_1 > \ldots > \mu_m = \ldots = \mu_t > \mu_{t+1} \geq \ldots \geq \mu_k. \tag{8}$$

From (5) and (8) we have

$$h_m > h_{m+1} \geq \mu_{m+1} = \mu_m = \mu_{t-1} = \mu_t \geq h_{n-k+t-1} > h_{n-k+t}. \tag{9}$$

Furthermore,

$$\begin{aligned} E_r(\mu) &= \mu_m E_{r-1}(\bar{\mu}_m) + E_r(\bar{\mu}_m) \\ &= \mu_t E_{r-1}(\bar{\mu}_t) + E_r(\bar{\mu}_t) \end{aligned} \tag{10}$$

where $E_q(\bar{\mu}_j)$ means $E_q(\mu_1, \ldots, \mu_{j-1}, \mu_{j+1}, \ldots, \mu_k)$. (If $r = k$, $E_r(\bar{\mu}_j) = 0$.) Now $E_{r-1}(\bar{\mu}_m) = E_{r-1}(\bar{\mu}_t) = 0$. For, if $E_r(\bar{\mu}_m) > 0$, then for $\mu' = (\mu_1, \ldots, \mu_m + \delta, \ldots, \mu_k)$,

$$E_r(\mu') = (\mu_m + \delta)\, E_{r-1}(\bar{\mu}_m) + E_r(\bar{\mu}_m) > E_r(\bar{\mu})$$

for $\delta > 0$, and, by (8) and (9), $\mu' \in R_k(h)$ for δ sufficiently small. This contradicts (6). Similarly, if $E_{r-1}(\bar{\mu}_t) < 0$, $E_r(\mu'') > E_r(\mu)$ for $\mu'' = (\mu_1, \ldots, \mu_t - \delta, \ldots, \mu_k)$. Hence $E_r(\mu) = E_r(\bar{\mu}_m)$ is independent of μ_m. Set $\nu_j = \mu_j$ for $j \neq m$, and choose $\nu_m > \mu_m$ so that $\nu_m < h_m$ and $\nu_m < \nu_{m-1}$ (if $m > 1$). Then $\nu \in T_{k,m+1}(h)$.

LEMMA 2. *Under the hypotheses of Lemma* 1,

$$\gamma = \max_{0 \leq s \leq k} E_r(h_1, \ldots, h_s, h_{n-k+s+1}, \ldots, h_n). \tag{11}$$

Proof. Since the lemma is obviously true when $r = 1$, and also when $k = n$, suppose that $2 \leq r \leq k < n$. By Lemma 1, $T_{kk}(h)$ is not empty. Let $S_{kq}(h)$, $1 \leq q \leq k$, be the set of those $\lambda \in T_{kk}(h)$ for which $\lambda_j = h_j$, $j = 1, \ldots, q$; and let $S_{k0}(h)$ be the set of $\lambda \in T_{kk}(h)$ for which $\lambda_1 < h_1$. Let s be the largest integer such that $S_{ks}(h)$ is not empty. If $s = k$, there is nothing to prove. Otherwise let $\mu \in S_{ks}(h)$. Then

$$\mu_j = h_{n-k+j},\ j = s + 1, \ldots, k;$$

for, if not, we shall show that there exists $\nu \in S_{k,s+1}(h)$, contradicting the choice of s.

Let t be the least integer greater than s for which $\mu_t > h_{n-k+t}$. If $t = s + 1$, $h_t > \mu_t$ by the maximality of s; while if $t > s + 1$

$$h_t \geq h_{n-k+t-1} = \mu_{t-1} > \mu_t.$$

Thus

$$h_t > \mu_t > h_{n-k+t}.$$

It follows that $E_{r-1}(\bar{\mu}_t) = 0$, since otherwise we could vary μ_t up or down to increase $E_r(\mu)$ (see (10)) while keeping μ in $T_{kk}(h)$.

Thus

$$E_r(\mu) = E_r(\bar{\mu}_t). \tag{12}$$

Set

$$\begin{aligned} \nu_j &= \mu_j,\ j = 1, \ldots, s, && \text{(if } s > 0) \\ \nu_{s+1} &= h_{s+1}, && \\ \nu_j &= \mu_{j-1},\ j = s+2, \ldots, t, && \text{(if } t > s+1) \\ \nu_j &= \mu_j,\ j = t+1, \ldots, k, && \text{(if } k > t). \end{aligned}$$

In effect, μ_t is replaced by h_{s+1}, and the resulting μ_j's are re-indexed to restore the ordering. By (12), $E_r(\nu) = E_r(\mu)$. It is then a straightforward matter to verify that $\nu \in S_{k,s+1}(h)$. This completes the proof of the lemma.

We are now in a position to complete the proof of the theorem. If the eigenvalues of H are distinct, then for o.n. $x_1, \ldots, x_k$,

$$\begin{aligned} g(x_1, \ldots, x_k) &\leqslant \max_{\lambda \in R_k(h)} E_r(\lambda) \\ &= E_r(h_1, \ldots, h_s, h_{n-k+s+1}, \ldots, h_n). \end{aligned}$$

for some s, $0 \leqslant s \leqslant k$. Now g attains this value for o.n. eigenvectors $y_1, \ldots, y_k$ corresponding to $h_1, \ldots, h_s, h_{n-k+s+1}, \ldots, h_n$, respectively. Thus

$$\max g = \max_{0 \leqslant s \leqslant k} E_r(h_1, \ldots, h_s, h_{n-k+s+1}, \ldots, h_n).$$

A similar result holds for the minimum. That these results remain valid when the eigenvalues of H are not all different follows by a continuity argument.

References

1. R. Courant and D. Hilbert, *Methods of mathematical physics*, vol. 1 (New York, 1953).
2. Ky Fan, *On a theorem of Weyl concerning eigenvalues of linear transformations, I*; Proc. N.A.S. (U.S.A.), *35* (1949), 652–5.
3. M. Marcus and J. L. McGregor, *Extremal properties of Hermitian matrices*, Can. J. Math., *8* (1956), 524–31.

The University of British Columbia

LINEAR TRANSFORMATIONS ON ALGEBRAS OF MATRICES: THE INVARIANCE OF THE ELEMENTARY SYMMETRIC FUNCTIONS

MARVIN MARCUS AND ROGER PURVES

1. Introduction. In this paper we examine the structure of certain linear transformations T on the algebra of n-square matrices M_n into itself. In particular if $A \in M_n$ let $E_r(A)$ be the rth elementary symmetric function of the eigenvalues of A. Our main result states that *if* $4 \leq r \leq n-1$ *and* $E_r(T(A)) = E_r(A)$ *for* $A \in M_n$ *then* T *is essentially* (*modulo taking the transpose and multiplying by a constant*) *a similarity transformation:*

$$T: A \rightarrow SAS^{-1}.$$

No such result as this is true for $r = 1, 2$ and we shall exhibit certain classes of counterexamples. These counterexamples fail to work for $r = 3$ and the structure of those T such that $E_3(T(A)) = E_3(A)$ for all $A \in M_n$ is unknown to us. In **(1)** it is established that those T which preserve the rank (determinant) of every matrix in M_n are essentially of the form $T: A \rightarrow PAQ$ where P and Q are non-singular, (PQ is unimodular). In the first part of what follows, we shall improve this result by requiring only that T preserves non-singularity. We remark that ***in general we do not assume that*** T ***is multiplicative or anti-multiplicative anywhere in the paper.***

We shall collect here the notation to be used throughout. For $A \in M_n$ let A' = transpose of A, $\rho(A)$ = rank of A, $\mathrm{tr}(A)$ = trace of A, A_{ij} = the element in position (i, j) of A, O_n = the n-square zero matrix, and E_{ij} = the n-square matrix with 1 at position (i, j), 0 elsewhere. In addition if $A \in M_p$ and $B \in M_q$ we define $A \oplus B \in M_{p+q}$ to be the direct sum of A and B. If $1 \leq p \leq n$ then Q_{pn} will be the set of all sequences of p-tuples $\omega = (i_1, \ldots, i_p)$ where $1 \leq i_1 < i_2 < \ldots < i_p \leq n$. A transformation T: $M_n \rightarrow M_n$ will be called a direct product if there exists a scalar c and fixed U and V in M_n such that

$$T(A) = cUAV$$

or

$$T(A) = cUA'V$$

for all $A \in M_n$. This is motivated by the fact that the mapping $T: A \rightarrow UAV$ has a matrix representation $V' \times U$, the direct product of V' and U, with a

Received October 17, 1958. The work of the first author was supported by United States National Science Foundation Grant NSF-G5416. The work of the second author was supported by the National Research Council of Canada.

proper choice of co-ordinate system for M_n. We remark that the mapping $T: A \to A'$ cannot be accomplished by pre- and post-multiplication by fixed matrices U and V for all A. We shall also denote by e.v.(A) the set of all n eigenvalues of A counting multiplicities.

2. Linear maps of GL_n into itself. As usual, GL_n is the group of n-square non-singular matrices in M_n. We shall determine all T such that $T(GL_n) \subseteq GL_n$.

LEMMA 2.1. *If $0 \neq A \in M_n$ then A is similar to a matrix B with $B_{ii} \neq 0$, $i = 1, \ldots, n$.*

Proof. We may assume A is in Jordan form. It is known in general that A is similar to a matrix with $\mathrm{tr}(A)/n$ in position (i, i), $i = 1, \ldots, n$. Hence we may assume $\mathrm{tr}(A) = 0$. If $A = E_{12}$ let u_1 be the vector with all entries 1 and let u_2 be the vector with first entry $1 - n$ and the remaining entries 1. Normalize u_1 and u_2 and let $u_3, \ldots, u_n$ be a completion to an orthonormal basis. Let U be the orthogonal matrix with u_i as column i. Then the (i, i) entry of $UE_{12}U'$ is $u_{i1}u_{i2} \neq 0$. The proof is now completed by induction on n. If $A \in M_{n+1}$ is in Jordan form with zero trace we consider first the case that A is diagonal. Since $A \neq 0$ we can assume $A_{11} \neq 0$ and moreover the matrix $C \in M_n$ obtained by deleting row and column 1 of A is not 0_n. By induction choose $V \in M_n$ such that $(VCV^{-1})_{ii} \neq 0$ for $i = 1, \ldots, n$. Then

$$(1 \oplus V)\,\mathrm{A}(1 \oplus V^{-1}) = \mathrm{A}_{11} \oplus VCV^{-1}$$

has all non-zero diagonal elements. If A is not diagonal we can clearly assume $A_{12} = 1$ and the submatrix C above is not 0_n. As before we select $V \in M_n$ such that

$$P = (1 \oplus V)\, A(1 \oplus V^{-1})$$

has all non-zero entries on the diagonal with the possible exception of P_{11}. If $P_{11} = 0$ and b_{11} is the $(1, 1)$ entry of VCV^{-1} then select $U \in M_2$ such that

$$U\begin{pmatrix} 0 & * \\ 0 & b_{11} \end{pmatrix} U^{-1}$$

has non-zero diagonal entries. Then

$$B = (U \oplus I_{n-1})\, P(U^{-1} \oplus I_{n-1})$$

is the required matrix.

LEMMA 2.2. *If $0 \neq A \in M_n$ then there is a $Z \in M_n$ such that*

$$\text{e.v.}(A + Z) \cap \text{e.v.}(Z) = 0.$$

Proof. By Lemma 2.1. choose $P \in M_n$ such that $(P^{-1}AP)_{ii \neq 0}$ for $i = 1, \ldots, n$. Let X be defined as follows:

$$X_{ii} = 1, \qquad i = 1, \ldots, n$$
$$X_{ij} = -(P^{-1}AP)_{ij}, \qquad i > j$$
$$X_{ij} = 0, \qquad i < j.$$

Then X has all n eigenvalues 1 and $P^{-1}AP + X$ has eigenvalues $1 + (P^{-1}AP)_{ii}$ $i = 1, \ldots, n$ none of which are 1. Then $Z = PXP^{-1}$ has the required property.

LEMMA 2.3. *If $T(GL_n) \subseteq GL_n$ then T is non-singular.*

Proof. We have that if

$$\det (xI_n - [T(I_n)]^{-1} T(A)) = 0$$

for some x then

$$\det (xI_n - A) = 0$$

for that x. In other words the distinct elements of e.v. $([T(I)]^{-1}T(A))$ form a subset of the distinct eigenvalues of A. Now suppose $0 \neq A \in M_n$ and $T(A) = 0$. Choose $Z \in M_n$ by Lemma 2.2 such that

$$\text{e.v. } (Z) \cap \text{e.v.}(A + Z) = 0$$

Then

$$[T(I_n)]^{-1} T(A + Z) = [T(I_n)]^{-1} T(Z)$$

and the distinct eigenvalues of $[T(I_n)]^{-1}T(Z)$ form a subset of the distinct eigenvalues of both $A + Z$ and Z. This shows that $A = 0$ if $T(A) = 0$ and T is non-singular.

LEMMA 2.4. *If $T(GL_n) \subseteq GL_n$ and $T(I_n) = I_n$ then* e.v. $(T(A)) =$ e.v.(A) *for all $A \in M_n$.*

Proof. As in the proof of Lemma 2.3, we know that if $T(A)$ has a set of n distinct eigenvalues then

$$\text{e.v.}(A) = \text{e.v.}(T(A)).$$

Since T^{-1} exists we can say that if B has n distinct eigenvalues then

$$\text{e.v.}(B) = \text{e.v.}(T^{-1}(B)).$$

If $T(A)$ has multiple eigenvalues choose a sequence B_j converging to $T(A)$ such that B_j has distinct eigenvalues. The proof is completed using the fact that the eigenvalues depend continuously on the elements.

THEOREM 2.1. *If $T(GL_n) \subseteq GL_n$ then there exist U and V in GL_n such that either*

$$T : A \to UAV \text{ for all } A \in M_n$$

or

$$T : A \to UA'V \text{ for all } A \in M_n.$$

Proof. By Lemma 2.4 the map

$$\phi : A \to [T(I_n)]^{-1}T(A)$$

satisfies

$$\text{e.v.}\ (\phi(A)) = \text{e.v.}(A)$$

for all $A \in M_n$. But by (**1**: Theorem 2),

$$\phi(A) = UAU^{-1}$$

or

$$\phi(A) = UA'U^{-1}.$$

Multiplication on the left by $T(I_n)$ completes the proof.

3. Linear maps preserving the symmetric functions. We now determine the structure of those linear T on M_n to M_n such that for each $A \in M_n$

$$E_r(A) = E_r(T(A)).$$

For each r let the class of all such T be denoted by $\mathfrak{A}_r$. It is clear that if T, $S \in \mathfrak{A}_r$ then $TS \in \mathfrak{A}_r$. Also if $T \in \mathfrak{A}_r$ and T^{-1} exists then $T^{-1} \in \mathfrak{A}_r$; for since any B is in the range of T we have

$$E_r(B) = E_r(TT^{-1}(B)) = E_r(T^{-1}(B)).$$

Our first result shows that $\mathfrak{A}_r$ is actually a multiplicative group for $r > 2$.

LEMMA 3.1. *If $r > 2$ and $T \in \mathfrak{A}_r$ then T^{-1} exists. Thus $\mathfrak{A}_r$ is a multiplicative group for $r > 2$.*

Proof. Suppose $T(A) = O_n$ and $A \neq O_n$. Then

$$E_r(A + X) = E_r(T(A + X)) = E_r(T(X)) = E_r(X)$$

for any $X \in M_n$. By Lemma 2.1 there exists $P \in GL_n$ such that $(P^{-1}AP)_{ii} \neq 0$ for $i = 1, \ldots, n$.

Define $X \in M_n$ as follows:

$$\begin{array}{ll} X_{ii} = x & i = 1, \ldots, r-1 \\ X_{ii} = 0 & i = r, \ldots, n \\ X_{ij} = 0 & i < j \\ X_{ij} = -(P^{-1}AP)_{ij} & i > j. \end{array}$$

Then

$$f_r(x) = E_r(P^{-1}AP + X) = E_r(A + PXP^{-1}) = E_r(PXP^{-1}) = E_r(X) = 0.$$

Thus the coefficient of x^{r-1} in the polynomial $f_r(x)$ must be 0. This means that the sum of the last $n - r + 1$ entries on the main diagonal of $P^{-1}AP$ is 0. Similarly we can show that the sum of any $n - r + 1$ is 0. But since $r > 2$, $n - r + 1 < n$ and it is clear that $(P^{-1}AP)_{ii} = 0$ $(i = 1, \ldots, n)$. This completes the proof.

LEMMA 3.2. *If* $A \in M_n$ *and* $A \neq 0$ *then*

$$\deg \det (xA + B) \leq 1 \text{ for all } B \in M_n$$

if, and only if, $\rho(A) = 1$.

Proof. We can clearly assume that A is in Jordan canonical form and the "if" part of the result is obvious.

In the other direction we show first that A has at most one non-zero eigenvalue. Suppose

$$\lambda_{i_1}, \ldots, \lambda_{i_k}$$

are the non-zero eigenvalues of A in positions (i_t, i_t), $t = 1, \ldots, k$. Let B be a diagonal matrix with 0 at positions (i_t, i_t) $t = 1, \ldots, k$ and 1 elsewhere on the main diagonal. Then

$$\deg \det (x\,A + B) = k = 1.$$

Suppose now that A has the single non-zero eigenvalue λ which we may assume is in position $(1, 1)$. To show that $\rho(A) = 1$ it will suffice to show that the elements along the superdiagonal of A are all 0. This is clear for $n = 2$. If $n > 2$ let α be the largest integer such that there is a 1 at position $(\alpha, \alpha + 1)$ of A. Define B as follows:

$$\begin{aligned} &B_{ii} = 0 && i = \alpha, \alpha + 1 \\ &B_{ii} = 1 && i \neq \alpha, \alpha + 1 \\ &B_{\alpha+1,\alpha} = 1 && \\ &B_{ij} = 0 && \text{elsewhere.} \end{aligned}$$

Then

$$\det (xA + B) = -\lambda x^2 - x.$$

Thus there must be a 0 at $(\alpha, \alpha + 1)$ and a repetition of this procedure shows that there are no 1's along the superdiagonal when $\lambda \neq 0$.

Now assume that $\lambda = 0$ and that the $(1, 2)$ entry of A is 1. Define α as above and if $\alpha > 2$ define B as follows:

$$\begin{aligned} &B_{ii} = 0 && i = 1, 2, \alpha, \alpha + 1 \\ &B_{ii} = 1 && \text{elsewhere on the main diagonal} \\ &B_{21} = 1 && \\ &B_{ij} = 0 && \text{elsewhere off the main diagonal.} \end{aligned}$$

Then

$$\det (xA + B) = x^2.$$

In this way all elements $(i, i + 1)$ for $2 < i < n - 1$ are shown to be 0. To settle position $(2, 3)$ use the test matrix

$$B = E_{31} \oplus I_{n-3}.$$

for $E_{31} \in M_3$. This completes the proof.

LEMMA 3.3. *If $3 \leq r < n$ and $A \in M_n$, $A \neq O_n$ then the condition*

$$\deg E_r\,(xA + B) \leq 1$$

for all $B \in M_n$ implies that A has at most one non-zero eigenvalue.

Proof. We can again assume A is in Jordan canonical form with eigenvalues $\lambda_1, \ldots, \lambda_n$. Let $z_1, \ldots, z_n$ be indeterminates and let B be the diagonal matrix with $B_{ii} = z_i$ $i = 1, \ldots, n$. Then

$$E_r(xA + B) = \sum_{\omega=(i_1,\ldots,i_r)\in Q_{r,n}} \prod_{k=1}^{r} (x\lambda_{i_k} + z_{i_k})$$

$$= \sum_{t=0}^{r} \Big(\sum_{\omega\in Q_{r,n}} \sum_{s_t\subseteq\omega} \prod_{\alpha\in s_t} \lambda_\alpha \prod_{\beta\in\omega-s_t} z_\beta\Big) x^t$$

where

$$\sum_{s_t\subseteq\omega}$$

means the sum over all subsets s_t of ω with t members and

$$\prod_{\beta\in\omega-s_t}$$

means the product over those elements of ω not in s_t. Hence for $t \geq 2$ we have that the coefficient of x^t in the above sum must be 0 for any choice of $z_1, \ldots, z_n$. From this it is not difficult to show that the tth elementary symmetric function of any $n - r + t$ of the λ_j is 0. Choosing $t = 2$ we have that if all the λ_j are equal they must all be 0. Assume then that for some μ, σ, $\lambda_\sigma \neq \lambda_\mu$. Since $r \geq 3$ we have that $k = n - r + 2 < n$. Let

$$\lambda_{i_1}, \ldots, \lambda_{i_{k-1}}$$

be a choice of $k - 1$ of the eigenvalues with $i_j \neq \sigma, \mu$ for $j = 1, \ldots, k - 1$. Then

$$0 = E_2\,(\lambda_\sigma, \lambda_{i_1}, \ldots, \lambda_{i_{k-1}}) = \lambda_\sigma E_1(\lambda_{i_1}, \ldots, \lambda_{i_{k-1}}) + E_2\,(\lambda_{i_1}, \ldots, \lambda_{i_{k-1}})$$

and a similar relation holds for λ_μ.
We then have

$$(\lambda_\sigma - \lambda_\mu)\, E_1\,(\lambda_{i_1}, \ldots, \lambda_{i_{k-1}}) = 0.$$

If $r > 3$ then $k - 1 < n - 2$ and this last relation implies that $\lambda_i = 0$ for $i \neq \sigma, \mu$. In this case

$$\lambda_\sigma\lambda_\mu = 0$$

and A has at most one non-zero eigenvalue. To settle the case $r = 3$ let $E_t(\hat{\lambda}_j)$ denote the tth elementary symmetric function of all the λ_i for $i \neq j$. We first note that

$$E_2(\lambda_1, \ldots, \lambda_n) = \lambda_j E_1(\hat{\lambda}_j) + E_2(\hat{\lambda}_j) = \lambda_j E_1(\hat{\lambda}_j).$$

Summing on j we have

$$n\, E_2(\lambda_1, \ldots, \lambda_n) = 2\, E_2(\lambda_1, \ldots, \lambda_n) = 0.$$

Thus

$$\lambda_j E_1(\hat{\lambda}_j) = 0.$$

Setting

$$s = \sum_{j=1}^{n} \lambda_j$$

we have

$$\lambda_j^2 = \lambda_j s, \; \lambda_j(\lambda_j - s) = 0$$

and thus the non-zero eigenvalues of A are all equal to s. This completes $r = 3$.

LEMMA 3.4. *Assume* $4 \leq r \leq n + 3$ *and let* $A \in M_{n+3}$, $A \neq O_n$. *Then*

$$\deg E_r(xA + B) \leq 1 \quad \textit{for all} \quad B \in M_{n+3}$$

if, and only if, $\rho(A) = 1$.

Proof. The "if" part of the theorem is clear. To prove the "only if" part we can assume A is in Jordan canonical form and proceed by induction on n. For $n = 1$ or $r = n + 3$ Lemma 3.2 gives the result. Thus assume $r < n + 4$ and by Lemma 3.3 we know that A has at most one non-zero eigenvalue λ which we can assume is in position (1, 1). Call the (2, 3) entry ϵ (either 1 or 0). Define B to be the matrix with 1 in position (3, 2) and $r - 3$ 1's in any of the diagonal positions (i, i) for $i > 3$, 0's elsewhere. Then

$$E_r(xA + B) = \lambda \epsilon x^2.$$

Consider first the situation in which $\lambda \neq 0$. Then $\epsilon = 0$ and row 2 and column 2 of A are both zero. If we restrict B to those matrices with row 2 and column 2 zero we can apply the induction hypothesis to conclude that the submatrix of A obtained by deleting row 2 and column 2 has rank 1. Thus $\rho(A) = 1$ as well. In case $\lambda = 0$ let ϵ_1 and ϵ_2 be the (1, 2) and $(n + 3, n + 4)$ entries of A respectively. Define B as follows:

$$\begin{aligned} &B_{21} = B_{n+4,n+3} = 1. && \\ &B_{ii} = 1, && 3 \leq i \leq r - 2 \\ &B_{ij} = 0 && \text{elsewhere.} \end{aligned}$$

Then

$$E_r(xA + B) = \epsilon_1 \epsilon_2 x^2$$

and we may assume without loss of generality that $\epsilon_2 = 0$. But then we can apply the induction argument as before to obtain $\rho(A) = 1$.

LEMMA 3.5. *If* $4 \leq r \leq n$ *and* $T \in \mathfrak{A}_r$ *and* $\rho(A) = 1$ *for* $A \in M_n$ *then* $\rho(T(A)) = 1$.

Proof. Consider the polynomial $f_r(x) = E_r(xT(A) + B)$. Since $T^{-1} \in \mathfrak{A}_r$

we have $f_r(x) = E_r(xA + T^{-1}(B))$. Since $\rho(A) = 1$, $\deg f_r(x) \leq 1$ for all B, and by Lemma 3.4 $\rho(T(A)) = 1$.

LEMMA 3.6. *If $4 \leq r \leq n$ and $T \in \mathfrak{A}_r$ then for every $A \in M_n$*

$$\rho(T(A)) = \rho(A).$$

Proof. Let $\rho(A) = k$ and select A_j $j = 1, \ldots, k$ such that $\rho(A_j) = 1$ and

$$A = \sum_{j=1}^{k} A_j.$$

Then by Lemmas 3.5 and 3.1

$$\rho(T(A)) \leq k = \rho(A) = \rho(T^{-1}(T(A)) \leq \rho(T(A)).$$

We are now in a position to prove our main result concerning the structure of $\mathfrak{A}_r$.

THEOREM 3.1. *If $4 \leq r \leq n - 1$ and $T \in \mathfrak{A}_r$ then there exist U and V in M_n such that either*

(i) $$T : A \to UAV \quad \textit{for all} \quad A \in M_n$$

or

(ii) $$T : A \to UA'V \quad \textit{for all } A \in M_n$$

where

(iii) $$UV = e^{i\phi}I_n,\ r\phi \equiv 0\ (2\pi).$$

Proof. The existence of U and V satisfying (i) and (ii) is an immediate consequence of Lemma 3.6 and Theorem 2.1. It is clear that it suffices to show that $E_r(PB) = E_r(B)$ for all $B \in M_n$ implies that $P = e^{i\phi}I_n$ with $r\phi \equiv 0(2\pi)$. Letting $C_r(B)$ denote the rth compound of B we have

$$\operatorname{tr} C_r(PB) = \operatorname{tr} C_r(B) \text{ for all } B \in M_n.$$

Hence

$$\operatorname{tr}\{[C_r(P) - I_{\binom{n}{r}}]\, C_r(B)\} = 0.$$

This implies immediately that

$$C_r(P) = I_{\binom{n}{r}}.$$

By the polar factorization theorem let $P = UH$ where U is unitary and H is positive definite Hermitian (p. d. h.). Then

$$C_r(U)C_r(H) = I_{\binom{n}{r}}$$

implies that $C_r(U)$ is both unitary and p. d. h. Hence every eigenvalue of $C_r(U)$ is 1 and this in turn implies that every eigenvalue of U is $e^{i\phi}$ for $r\phi \equiv 0\ (2\pi)$. Similarly we show $H = I_n$ and the result is at hand.

4. The structure of $\mathfrak{A}j$ for $j = 1, 2, 3$. At this point Theorem 3.1 together with the results in **(1)** completely settle the question of the structure of $\mathfrak{A}_r$ when $r > 4$. It is easy to construct singular $T \in \mathfrak{A}_1$ (map A into the diagonal matrix B with $B_{ii} = A_{ii}$). Thus not much can be said about $\mathfrak{A}_1$. In examining $\mathfrak{A}_2$ we are led to two kinds of counterexamples: (i) those transformations $S \in \mathfrak{A}_2$ which permute the entries of every $A \in M_n$ in some fixed way; (ii) those transformations $C \in \mathfrak{A}_2$ which map A into $K \circ A$ where $K \in M_n$ and $K \circ A$ is the Hadamard product of K and A ($(K \circ A)_{ij} = K_{ij}A_{ij}\, i, j = 1, \dots, n$). We shall show that there exist non-trivial examples of both types (i) and (ii) in $\mathfrak{A}_2$ but that no such examples exist in $\mathfrak{A}_3$. We remark here that Lemma 3.4 fails for $r = 3$; for take $A = E_{12} + E_{34} \in M_4$ and note that although $E_3\,(x\,A + B)$ is at most linear in x for $B \in M_4$, $\rho(A) = 2$. Thus there is no hope for proving Theorem 3.1 via Lemma 3.4 for $r = 3$.

Denote by S_r that subset of $\mathfrak{A}_r$ consisting of transformations that rearrange the elements of every $A \in M_n$ in some fixed way. Similarly, let H_r denote that subset of $\mathfrak{A}_r$ consisting of transformations of the type $A \to K \circ A$, $K \in M_n$.

THEOREM 4.1. *If $S \in S_2$ then $S = \sigma_1\,\sigma_2\,\sigma_3$ where*

(i) *σ_3 is a permutation of the main diagonal entries only.*

(ii) *σ_2 is a permutation of the set of pairs of entries symmetrically located across the main diagonal.*

(iii) *σ_1 interchanges symmetrically located entries.*

The proof of Theorem 4.1 is a straightforward enumeration of the possibilities for images under S of matrices of the types $E_{ii} + E_{jj}$, $i < j$ and $E_{ij} + E_{ji}$, $i < j$. We omit the details.

THEOREM 4.2. *No element of S_2 of the types* (i), (ii), (iii) *in* Theorem 4.1 *is a direct product except the identity map and the transpose map.*

Proof. This is done by showing that any map of the types $\sigma_1, \sigma_2, \sigma_3$ described in Theorem 4.1 maps some non-singular N into a singular matrix. First, suppose σ_3 maps the (j, j) entry into the (i_j, i_j) entry. Choose a permutation π of $1, \dots, n$ such that $\pi(j) = j$ and $\pi(i) \neq i$ for $i \neq j$. Let N be the permutation matrix corresponding to π and observe that $\sigma_3(N)$ is singular. Next, suppose σ_2 maps (i, j) and (j, i) into (k, l) and (l, k) respectively. Let

$$N = E_{ij} + E_{ji} + \sum_{t \neq i,j} E_{tt}$$

and note that N is non-singular and $\sigma_2(N)$ is singular. Next, suppose σ_1 interchanges (i, j) and (j, i) and leaves fixed (k, l) and (l, k). It is not difficult to exhibit non-singular $N \in M_3$ or M_4 for which $\sigma_1(N)$ is singular and we proceed to show that the examples in M_n for $n > 4$ can be reduced to one of the cases $n = 3$ or $n = 4$. Suppose first that none of the equalities: $i = k$, $i = l$, $j = k$, $j = l$ holds. Then set $N_1 = E_{ij} + E_{ji} + E_{kl} + E_{lk}$ and let the permutation π of $1, \dots, n$ be $(i\,j)\,(2\,i)$ with corresponding permutation matrix

P. Then $P\sigma(N_1)P' = E_{12} + E_{21} + E_{kl} + E_{lk}$. Similarly obtain a permutation matrix Q such that $QP\sigma_1(N_1)P'Q' = E_{12} + E_{21} + E_{34} + E_{43}$. We are then confronted essentially with the case $n = 4$. If any of the equalities $i = k$, $i = l$, $j = k$, $j = l$ holds we can reduce the situation to the case $n = 3$ by a similar device.

We may describe the structure of H_2 as follows:

THEOREM 4.3. *If* $C \in H_2$, $C : A \to K \circ A$ *then* $K_{ij} = (K_{ji})^{-1}$ *for* $i \neq j$ *and either* $K_{ii} = 1$ $(i = 1, \ldots, n)$ *or* $K_{ii} = -1$ *for* $i = 1, \ldots, n$.

We omit the proof which consists of a straightforward consideration of the possibilities for the 2-square sub-determinants of K.

We remark at this point that it seems plausible that $\mathfrak{A}_2$ is generated by taking only products of elements of S_2, H_2 and maps of the form $A \to PAP^{-1}$, $P \in GL_n$. We have been unable to prove this, however.

The situations for S_3 and H_3 are somewhat more involved but we shall use a sequence of lemmas to show that:

S_3 consists only of the identity map, the transpose map, and maps of the form $A \to PAP'$ for P a permutation matrix; H_3 consists only of the identity map and the map $A \to K \circ A = \theta DAD^{-1}$ where D is a diagonal matrix and θ is a cube root of 1. It is not known to us whether there exist other elements of $\mathfrak{A}_3$ which are not direct products.

LEMMA 4.1. *If* $A \in M_n$ *and* A *has* n *elements* 1, *the rest* 0, *then for* $n \geq r \geq 1$,

$$E_r(A) = \binom{n}{r}$$

if, and only if, $A = I_n$.

Proof. It is clear that since the rth order subdeterminants of A are integers that

$$E_r(A) \leq \operatorname{tr}\{[C_r(A)]\,[C_r(A)]'\}.$$

Hence

$$E_r(A) \leq \operatorname{tr} C_r(AA') = E_r(\alpha_1^2, \ldots, \alpha_n^2)$$

where $\alpha_j^2, j = 1, \ldots, n$ are the eigenvalues of AA'. If $\rho(A) = k$ and $k < r$ it is clear that

$$0 = E_r(A) < \binom{n}{r}.$$

Otherwise if $k \geq r$

$$\begin{aligned} E_r(A) &\leq E_r(\alpha_1^2, \ldots, \alpha_n^2) = E_r(\alpha_1^2, \ldots, \alpha_k^2) \\ &\leq \binom{k}{r} k^{-r}\,\{E_1(\alpha_1^2, \ldots, \alpha_k^2)\}^r \\ &= \binom{k}{r} k^{-r}\,\{\operatorname{tr}(AA')\}^r = \binom{k}{r} k^{-r} n^r. \end{aligned}$$

We consider two cases:

(i) $k = n$. Then A is a permutation matrix and all eigenvalues lie on the unit circle. Then it is easily seen that

$$E_r(A) = \binom{n}{r}$$

implies all the eigenvalues are equal and the only permutation matrix with this property is I_n.

(ii) $k < n$. We shall show this is impossible. If $k = 1$, then $r = 1$ and $E_1(A) = \text{tr } (A) = n$. But I_n is the only matrix satisfying this and this is a contradiction. On the other hand, if $k \geqslant 2$ then

$$E_r(A) \leqslant \binom{k}{r} k^{-r} n^r < \binom{n}{r} = E_r(A)$$

and the proof is complete.

Lemma 4.2. *If $S \in S_3$ and $n \geqslant 4$ then S either interchanges (i, j) and (j, i) for $i \neq j$ or leaves them fixed.*

Proof. Since

$$E_3(S(I_n)) = \binom{n}{3}$$

we have $S(I_n) = I_n$ by Lemma 4.1. Thus we may modify S to obtain

$$\sigma : A \rightarrow PS(A)P'$$

where $P \in M_n$ is such a permutation matrix that σ holds the main diagonal elements fixed. Now let

$$N_0 = 0_{n-2} \oplus J_2$$

where

$$J_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

We show first that $\sigma(N_0) = N_0$. If this were not the case we have two possible alternatives:

(i) $\sigma(N_0)$ has a 1 at some position (k, l) such that $k < 1$ and $(k, l) \neq (n - 1, n)$.

(ii) $\sigma(N_0)$ has a 1 at some position (k, l) such that $k > 1$ and $(k, l) \neq (n, n - 1)$.

In (i) let D be a diagonal matrix in M_{n-2} with 1 at (k, k) and $(n - 3)$ zero's elsewhere on the diagonal. Then

$$E_3(D \oplus J_2) = -1.$$

However $\sigma(D \oplus J_2)$ has at most two non-zero rows and hence

$$E_3(\sigma(D \oplus J_2)) = 0.$$

In a similar way we eliminate the alternative (ii). Hence σ either interchanges or leaves fixed the entries at $(n-1, n)$ and $(n, n-1)$. A similar argument for the other pairs of symmetrically located entries completes the proof.

LEMMA 4.3. *If $S \in S_r$, $r > 2$ and*

$$S : A \to UAV$$

or

$$S : A \to UA'V$$

then U and V are permutation matrices.

We omit the proof.

THEOREM 4.4. *If $S \in S_3$ and $n = p + 2$, $p > 1$ then either*

$$S : A \to PAP' \quad \textit{for all} \quad A \in M_n$$

or

$$S : A \to PA'P' \quad \textit{for all} \quad A \in M_n$$

where $P \in M_n$ is a permutation matrix.

Proof. The proof is by induction on the integer p. For $p = 1$ the result in **(1**, Theorem 2**)** shows that S is a direct product (modulo taking the transpose), and Lemma 4.3 combined with argument used in the latter part of the proof of Theorem 3.1 establishes that S has the above form. Now we modify S as in Lemma 4.2 to obtain $\sigma \in S_3$ where σ holds diagonal elements fixed. Assume the result for all integers up to $p > 1$. Then if $C \in M_{n-1} = M_{(p-1)+2}$ we have by Lemma 4.2 that

$$\sigma(0 \oplus C) = 0 \oplus \sigma(C)$$

and

$$\begin{aligned} E_3(\sigma(C)) = E_3(0 \oplus \sigma(C)) &= E_3(\sigma(0 \oplus C)) \\ = E_3(0 \oplus C) &= E_3(C). \end{aligned}$$

By the induction hypothesis and the fact that σ holds the diagonal elements fixed we see that if we consider σ as a mapping of $M_{n-1} \to M_{n-1}$ in the obvious way then

$$\sigma : C \to C \text{ for all } C \in M_{n-1}$$

or

$$\sigma : C \to C' \text{ for all } C \in M_{n-1}.$$

Now it is clear that if $A \in M_n = M_{p+2}$ and $C_i \in M_{n-1}$ is the principal submatrix obtained by deleting row and column i of A then the above argument shows that

$$\sigma(C_i) = C_i$$

or

$$\sigma(C_i) = C_i'.$$

Thus for each $A \in M_n$ it follows that

$$\sigma(A) = A$$

or

$$\sigma(A) = A',$$

and the proof is complete.

THEOREM 4.5. *If* $C \in H_3$ *then there exists* $D \in M_n$ *such that*

$$C : A \to \theta\, DAD^{-1} \quad \text{for all} \quad A \in M_n$$

where D *is a diagonal matrix and* $\theta^3 = 1$.

Proof. It suffices to show that there exist diagonal U and V in M_n such that $C(A) = UAV$ or $C(A) = UA'V$ for then it is clear that $U_{tt} = \theta^{-1}V_{tt}^{-1}$ for $i = 1, \ldots, n$ and $\theta^3 = 1$. Now for each $\omega \in Q_{3n}$ it is clear that we may consider C as a mapping of $M_3 \to M_3$ by restricting C to the principal submatrix of each $A \in M_n$ corresponding to the indices of ω. Call the restricted mapping $C_\omega : M_3 \to M_3$; and since C_ω preserves determinant it is a direct product:

$$C_\omega : A \to U_\omega A V_\omega \quad \text{for} \quad A \in M_3.$$

It is easy to check that U_ω and V_ω are diagonal by examining the images of $E_{tt} \in M_3$, $i = 1, 2, 3$ and using the fact that $C_\omega(A)$ is a Hadamard product. Thus on each 3-square principal submatrix C has the desired form. It will clearly suffice to show that $C : A \to K \circ A$ has the property $\rho(K) = 1$. For then K has the form $K_{ij} = a_i b_j$ $i, j = 1, \ldots, n$. We show that every 2-square submatrix of K is singular. Let $(\alpha_i\, \beta_j)$ denote the submatrix of K involving rows $\alpha_1\, \alpha_2$ and columns β_1, β_2. Suppose $\{\alpha_1, \alpha_2, \beta_1, \beta_2\}$ involves fewer than 4 distinct integers. Then it is clear that $(\alpha_i\, \beta_j)$ is a part of some principal 3-square submatrix whose row and column indices we will designate by

$$\theta = \{\gamma_1\gamma_2\gamma_3\}.$$

By the above argument C_θ has the form

$$C_\theta : A \to U_\theta A V_\theta;\ A \in M_3$$

where U_θ and V_θ are diagonal with diagonal elements u_1, u_2, u_3 and v_1, v_2, v_3 respectively. It follows that for some i_1, i_2, j_1, j_2 that

$$K_{\alpha_s\beta_t} = u_{i_s} v_{j_t}\ s, t = 1, 2$$

and hence that $(\alpha_i\, \beta_j)$ is singular. In case $\{\alpha_1, \alpha_2, \beta_1, \beta_2\}$ consists of 4 distinct integers we consider the two 3-square principal submatrices corresponding to

$$\mu = \{\alpha_1, \alpha_2, \beta_1\} \quad \text{and} \quad \sigma = \{\alpha_1, \alpha_2, \beta_2\}.$$

Again we see that

$$C_\mu : A \to U_\mu A V_\mu,\ A \in M_3$$
$$C_\sigma : A \to U_\sigma A V_\sigma,\ A \in M_3$$

where U_μ, U_ν, V_μ and V_ν are diagonal with main diagonals

$$(u_1, u_2, u_3),\ (u'_1, u'_2, u'_3),\ (v_1, v_2, v_3),\ (v'_1, v'_2, v'_3)$$

respectively. We then obtain for some i_1, j_1, i_2.

$$\begin{aligned} K_{\alpha_1\beta_1} &= u_{i_1}v_{j_1} & K_{\alpha_1\alpha_1} &= u_{i_1}v_{i_1} \\ K_{\alpha_2\beta_1} &= u_{i_2}v_{j_1} & K_{\alpha_2\alpha_1} &= u_{i_2}v_{i_1} \end{aligned}$$

and for some n_1, n_2, m_2,

$$\begin{aligned} K_{\alpha_1\beta_1} &= u'_{n_1}v'_{m_2} & K_{\alpha_1\alpha_1} &= u'_{n_1}v'_{n_1} \\ K_{\alpha_2\beta_2} &= u'_{n_2}v'_{m_2} & K_{\alpha_2\alpha_1} &= u'_{n_2}v'_{n_1}. \end{aligned}$$

From these equalities we see that

$$K_{\alpha_1\beta_1}/K_{\alpha_2\beta_1} = K_{\alpha_1\beta_2}/K_{\alpha_2\beta_2}$$

and again $(\alpha_i\, \beta_j)$ is singular.

Reference

1. M. Marcus and B. N. Moyls, *Linear transformations on algebras of matrices*, Can. J. Math., *11* (1959), 61–6.

University of British Columbia

PRIME DUAL IDEALS IN BOOLEAN ALGEBRAS

L. J. HEIDER

1. Introduction. Let $\mathfrak{B}$ denote an arbitrary Boolean algebra. Let Latin letters $a, b, \ldots$ denote general elements of $\mathfrak{B}$ while the symbols 0, 1 denote the special smallest and largest elements. Let Greek letters $\alpha, \beta, \ldots$ denote various prime dual ideals of elements of $\mathfrak{B}$. It is recalled that a prime dual ideal of $\mathfrak{B}$ is a proper subset of $\mathfrak{B}$ closed under finite intersections of its elements and maximal with respect to those properties. Every prime dual ideal includes the element 1 and for each element a of $\mathfrak{B}$ includes either a or $\bar{a}$ (complement of a in $\mathfrak{B}$) but not both. Occasional reference will be made to principal dual ideals of $\mathfrak{B}$. These are subsets of $\mathfrak{B}$ composed of all elements of $\mathfrak{B}$ majorizing some fixed non-zero element of $\mathfrak{B}$. Finally, let $X(\mathfrak{B})$ denote the collection of all prime dual ideals of $\mathfrak{B}$. Then, with the subsets $X(a) = [\alpha \in X(\mathfrak{B}) | a \in \alpha]$, $a \in \mathfrak{B}$, being used as a basis for open sets, the collection $X(\mathfrak{B})$ becomes (homeomorphic to) the Stone representation space for $\mathfrak{B}$.

The collection $X(\mathfrak{B})$, with its field of open-and-closed subsets, is primarily representative of the Boolean algebra $\mathfrak{B}$. Special field-related properties of particular algebras $\mathfrak{B}$ as, for example, the ability of $\mathfrak{B}$ to be represented as a quotient-field of sets, appear as special properties of the field $X(\mathfrak{B})$. However, the same collection $X(\mathfrak{B})$, with its compact, zero-dimensional, Hausdorff topology, may, with equal ease, be regarded as the Stone-Čech compactification space βY of a completely regular topological space Y. In this case, the algebra $\mathfrak{B}$ is provided by a basis of open-and-closed subsets of Y, and special properties of Y appear as special properties of $X(\mathfrak{B})$ and $\mathfrak{B}$.

In either case, it is the points of $X(\mathfrak{B})$ that matter. These points are not undefined terms, but complex structures, that is, prime dual ideals of a Boolean algebra $\mathfrak{B}$. Any prime dual ideal α of $\mathfrak{B}$ has the property that if a finite union element $\bigvee_{i=1}^{n} a_i$ of $\mathfrak{B}$ is in α, then some component element a_i of this union is likewise in α. This universal property of prime dual ideals may obviously be generalized. Let $\mathfrak{M}$ denote an infinite cardinal, and let I denote an index set of cardinality $\mathfrak{M}$. Assume that a union element $a_0 = \bigvee_{i \in I} a_i$ exists in $\mathfrak{B}$. In general, a prime dual ideal of $\mathfrak{B}$ containing a_0 may or may not contain a component element of this union.

This paper discusses the presence in $X(\mathfrak{B})$ of prime dual ideals that contain along with a union element $a_0 = \bigvee_{i \in I} a_i$ also a component element a_i of that union. The first result of this discussion is a unified theory of the use of $X(\mathfrak{B})$ in the representation of Boolean algebras $\mathfrak{B}$. Since the parts of this theory

Received December 16, 1957; in revised form October 15, 1958. This paper was prepared with the support of the National Science Foundation under Research Grant NSF-G3965.

have been developed by many authors, the present treatment is in outline form. The emphasis is on the unity of theory achieved by use of the above special property of prime dual ideals. The second result is a characterization of the Boolean algebras $\mathfrak{B}$ for which the spaces $X(\mathfrak{B})$ may be regarded as the Stone-Čech compactification spaces βY associated with three special types of completely regular spaces Y, namely, the P-, P'- and U-spaces of **(3, 4)**. These special spaces were introduced because of the interest of the algebraic features of their associated rings of real-valued continuous functions. Our interest arose from the fact that for each space Y of any of these types the corresponding space βY is zero-dimensional and thus homeomorphic to the representation space $X(\mathfrak{B})$ of a Boolean algebra $\mathfrak{B}$. In the cases of the P- and P'-spaces, the points of $\beta Y = X(\mathfrak{B})$ corresponding to points in Y involve intriguing properties of prime dual ideals.

2. Boolean algebras and fields of sets. Let $\mathfrak{M}$ denote an arbitrary cardinal number. Let the concepts of a field of sets, an $\mathfrak{M}$-field of sets and an $\mathfrak{M}$-complete Boolean algebra be understood in the usual sense. An $\mathfrak{M}$-complete Boolean algebra is called *$\mathfrak{M}$-representable* if it is isomorphic to an $\mathfrak{M}$-field of sets modulo an $\mathfrak{M}$-complete ideal of that field. An $\mathfrak{M}$-complete Boolean algebra $\mathfrak{B}$ is called *$\mathfrak{M}$-distributive* if

$$\bigvee_{i \in I} \bigwedge_{j \in J} a_{ij} = \bigvee_{h \in J^I} \bigwedge_{i \in J} a_{i,h(i)}$$

for each doubly-indexed family $\{a_{ij}\}$, $i \in I, j \in J$, of elements of $\mathfrak{B}$ for which the cardinalities $\bar{I}$, $\bar{J}$ of the index sets do not exceed $\mathfrak{M}$. Here J^I indicates the family of all maps h with domain I and range J.

For any element a_0 of a given Boolean algebra $\mathfrak{B}$ let $a_0 = \bigvee_{i \in I} a_i$, $\bar{I} < \mathfrak{M}$, be called an *$\mathfrak{M}$-representation* of the element a_0. Let

$$a_0 = \bigvee_{j \in J_i} a_{ij}, i \in I, \bar{I} < \mathfrak{M}, \bar{J}_i < \mathfrak{M},$$

be called an $\mathfrak{M}$-family of $\mathfrak{M}$-representations of a_0. With this terminology and these concepts at hand, the principal parts of the theory may be presented in three statements.

(A) The Boolean algebras that are isomorphic to $\mathfrak{M}$-fields of sets are the $\mathfrak{M}$-complete algebras that have for every non-zero element a prime dual ideal that contains a component of each $\mathfrak{M}$-representation of that element.

(B) The $\mathfrak{M}$-complete and $\mathfrak{M}$-distributive Boolean algebras are exactly those $\mathfrak{M}$-complete algebras that have for each non-zero element and for each $\mathfrak{M}$-family of $\mathfrak{M}$-representations of that element a principal dual ideal containing a component of each member of that family.

(C) The $\mathfrak{M}$-complete and $\mathfrak{M}$-representable Boolean algebras are exactly those $\mathfrak{M}$-complete algebras that have for each non-zero element and for each $\mathfrak{M}$-family of $\mathfrak{M}$-representations of that element a prime dual ideal containing a component of each member of that family.

These statements are made without proof. Their intended value lies in the unified treatment of diverse subjects that they provide. Statement (A) is an observation of Sikorski **(10)** in dual form. Enomoto's theorems **(2)** regarding $\mathfrak{M}$-fields of sets *in the wider sense* involve but slight rephrasing of this statement. Statement (B) is well known **(9, 11)**, but attention is here called to the position of $\mathfrak{M}$-distributive algebras midway between $\mathfrak{M}$-fields of sets and quotients of such fields by $\mathfrak{M}$-complete ideals. Statement (C) was suggested by work of Chang **(1)**, but is new at least in its simplicity.

An apparent addition to the existing literature on the subject matter of statement (C) may well be made here. Let $\mathfrak{B}$ be an $\mathfrak{M}$-complete Boolean algebra with representation space $X(\mathfrak{B})$. Let $\mathfrak{F}(\mathfrak{B})$ denote the $\mathfrak{M}$-field of subsets of $X(\mathfrak{B})$ generated by the subsets of $X(\mathfrak{B})$ of the type $X(a) = [\alpha \in X(\mathfrak{B}) \,|\, a \in \alpha]$, $a \in \mathfrak{B}$. Let an element of $\mathfrak{F}(\mathfrak{B})$ of the form $\bigcap_{j \in J} X(a_j)$ with $\bar{J} < \mathfrak{M}$ and $\bigwedge_{j \in J} a_j = 0$ in $\mathfrak{B}$ be called an $\mathfrak{M}$-nowhere dense subset of $X(\mathfrak{B})$. Let $\mathfrak{J}(\mathfrak{B})$ denote the $\mathfrak{M}$-complete ideal in $\mathfrak{F}(\mathfrak{B})$ generated by these $\mathfrak{M}$-nowhere dense subsets. Attention is now called to the fact that, for each $\mathfrak{M}$-complete and $\mathfrak{M}$-representable Boolean algebra $\mathfrak{B}$, the quotient $\mathfrak{F}(\mathfrak{B})/\mathfrak{J}(\mathfrak{B})$ is a specific example of an isomorphic representation of $\mathfrak{B}$ as the quotient of a $\mathfrak{M}$-field of sets modulo an $\mathfrak{M}$-complete ideal.

3. Fields of sets and topological spaces. The concept of a field of sets stands midway between that of a Boolean algebra and that of a topological space with a basis of open-and-closed subsets. Let $\mathfrak{M}$ denote an arbitrary cardinal number. Let $\mathfrak{F}(X)$ be an $\mathfrak{M}$-field of subsets of a set X. It will be assumed that $\mathfrak{F}(X)$ is *reduced*, that is, for $p \neq q$ in X there is an element O of $\mathfrak{F}(X)$ with $p \in O$ and $q \notin O$. Let $(X, \mathfrak{T})$ denote the set X as under the topology $\mathfrak{T}$ obtained by using the subsets of X in $\mathfrak{F}(X)$ as a basis for open sets. Any subset of X in $\mathfrak{F}(X)$ is open-and-closed in $(X, \mathfrak{T})$. However, there might be subsets of X not in $\mathfrak{F}(X)$ that are open-and-closed in $(X, \mathfrak{T})$. They would be of the form

$$A = \bigcup_{i \in I} O_i = \bigcap_{j \in J} O_j$$

where the index sets I, J are arbitrary and each O_i and O_j is an element of $\mathfrak{F}(X)$. This introduction of alien open-and-closed subsets will be undesirable for our purpose. Hence, a reduced $\mathfrak{M}$-field of sets $\mathfrak{F}(X)$ will be called *union-intersection closed* if every subset A of X as described above is an element of $\mathfrak{F}(X)$. With each reduced, $\mathfrak{M}$-field there is associated a minimal, reduced, union-intersection closed, $\mathfrak{M}$-field including the given field. It consists of all subsets A as described above.

We now turn to the very special topological spaces described in **(3, 4, 8)**. As usual, for any topological space Y, $C(Y)$ will denote the collection of all real-valued functions, defined and continuous on Y. For each element f of $C(Y)$, let $P(f) = [p \in Y \,|\, f(p) > 0]$ and $Z(f) = [p \in Y \,|\, f(p) = 0]$. Let βY and υY denote, respectively, the Stone-Čech compactification space and the

Hewitt Q-space associated with a completely regular space Y. The first special completely regular spaces to be considered are the P-spaces.

The P-spaces may be characterized in a number of different ways (**3**, Theorem 5.3). For one thing, a completely regular space Y is a P-space if, and only if, every countable intersection of open sets of Y is itself open in Y. From this it follows that each P-space Y is a zero-dimensional Hausdorff space in which each countable intersection of open-and-closed subsets is open-and-closed. Hence, dually, in a P-space any countable union of open-and-closed subsets is likewise open-and-closed. Thus, if Y is a P-space and $\mathfrak{F}(Y)$ is the field of open-and-closed subsets of Y, then $\mathfrak{F}(Y)$ is a reduced, union-intersection closed, σ-field of sets in the sense explained above.

Conversely, let $\mathfrak{F}(Y)$ be a reduced, union-intersection closed, σ-field of sets. Use the subsets of Y in $\mathfrak{F}(Y)$ as the basis of a topology $\mathfrak{T}$ on Y, and let $(Y, \mathfrak{T})$ denote Y with this topology.

THEOREM 3.1. *If $\mathfrak{F}(Y)$ is a reduced, union-intersection closed, σ-field of sets, then $(Y, \mathfrak{T})$ is a P-space and every P-space may be thus described.*

Proof. With $\mathfrak{F}(Y)$ and $(Y, \mathfrak{T})$ as described, it is obvious that $(Y, \mathfrak{T})$ is a zero-dimensional Hausdorff space and thus completely regular. Consider, moreover, the intersection $\bigcap U_n$ of a countable family $\{U_n\}$ of sets open in $(Y, \mathfrak{T})$. If p_0 is a point of Y in this intersection, then there exists a family $\{O_n\}$ of sets in $\mathfrak{F}(Y)$ with $p_0 \in O_n < U_n$ for each n. Hence, with $\mathfrak{F}(Y)$ a σ-field, there exists an element O_0 of $\mathfrak{F}(Y)$ with $p_0 \in O_0 \subseteq U_n$ for each n. Thus any countable intersection of open subsets of $(Y, \mathfrak{T})$ is open, so that $(Y, \mathfrak{T})$ is a P-space.

If, conversely, one begins with a P-space Y and then forms $\mathfrak{F}(Y)$ and $(Y, \mathfrak{T})$ as described, clearly $(Y, \mathfrak{T})$ is homeomorphic to Y.

With the P-spaces thus firmly linked to reduced, union-intersection closed, σ-fields of sets, attention is turned elsewhere for the moment. First, two additional facts (**3**, Theorem 5.3, (2) and (3)) concerning P-spaces are needed: if Y is a P-space, so likewise is νY; if Y is a P-space, then the zero-set $Z(f)$ is open-and-closed in Y for each element f of $C(Y)$.

Now, for any completely regular space Y and for any point p_0 in Y, let p_0 be called a P-point of Y if for each element f of $C(Y)$ there exists a neighbourhood U of p_0 in Y such that $f(p) = f(p_0)$ for each point p in U. Then, from the facts cited just above, it follows that for any P-space Y each point of υY is a P-point of υY. Next consider $\beta Y = \beta(\upsilon Y)$. It is rather obvious that each P-point of υY as imbedded in βY becomes a P-point of βY. On the other hand, no point $\bar{p}$ of $\beta Y - \upsilon Y$ as in βY is a P-point of βY. Thus, for each point $\bar{p}$ of this type, there is an element f of $C(\beta Y)$ with $f(\bar{p}) = 0$ while $f(p) > 0$ for all points p of υY (**5**, Example 2.3). This, of course, excludes the local constancy of f at $\bar{p}$ since the points of νY are dense in βY. Thus, for any P-space Y, the points of υY as imbedded in βY are identified with the P-points of βY.

The fact that each zero-set $Z(f)$ associated with a P-space Y is open-and-closed in Y indicates that for such spaces the sets $\bar{P}(f)$ are likewise open-and-closed in Y. Thence it follows (**4**, Theorem 8.3) that for any P-space Y the lattice $C(Y)$ is conditionally countably complete, so that $\beta Y = X(\mathfrak{B})$ where $\mathfrak{B}$ is a σ-complete Boolean algebra (**12**). This algebra may, of course, be identified with the Boolean algebra of all open-and-closed subsets of βY or, equivalently, of vY or even of Y itself.

4. P-spaces and Boolean algebras. Interest now turns to the P-points of a space $X(\mathfrak{B})$ where $\mathfrak{B}$ is a σ-complete Boolean algebra. Each point of $X(\mathfrak{B})$ is a prime dual ideal of $\mathfrak{B}$. Let α be such an ideal while $\mathfrak{M}$ is a cardinal number and I is an index set with $\bar{I} < \mathfrak{M}$. We introduce two conditions:

$(I - \mathfrak{M})$ If $\vee_{\iota\epsilon I}\, a_\iota$ exists and is in α, $\bar{I} < \mathfrak{M}$, then some a_ι is in α.
$(II - \mathfrak{M})$ If $\{a_\iota, \iota \in I\} \subseteq \alpha$, $\bar{I} < \mathfrak{M}$, then $\wedge_{\iota\epsilon I}\, a_\iota$ exists and is non-zero.

For $\mathfrak{M}$-complete Boolean algebras the two conditions are equivalent. For any Boolean algebra, if condition $II - \mathfrak{M}$ is satisfied with respect to a particular prime dual ideal, then condition $I - \mathfrak{M}$ is satisfied also.

A Boolean algebra $\mathfrak{B}$ will be called a $\mathfrak{B}(II - \mathfrak{M}, D)$ algebra if the prime dual ideals of $\mathfrak{B}$ satisfying condition $II - \mathfrak{M}$ are dense in $X(\mathfrak{B})$ or, equivalently, each element of $\mathfrak{B}$ is contained in a prime dual ideal of $\mathfrak{B}$ satisfying this condition. For each $\mathfrak{B}(II - \mathfrak{M}, D)$ algebra $\mathfrak{B}$, let D also denote the subspace of $X(\mathfrak{B})$ consisting of all points (prime dual ideals) satisfying condition $II - \mathfrak{M}$. A similar definition and notation can be used for $\mathfrak{B}(I - \mathfrak{M}, D)$ algebras. Although reference is made to an arbitrary cardinal number $\mathfrak{M}$, interest centers on the first infinite cardinal number $\aleph_0 = \sigma$. Two lemmas are now in order.

LEMMA 4.2. *Every $\mathfrak{B}(II - \mathfrak{M}, D)$ Boolean algebra is $\mathfrak{M}$-complete.*

LEMMA 4.3. *For any $\mathfrak{B}(II - \sigma, D)$ Boolean algebra $\mathfrak{B}$, the P-points of the space $X(\mathfrak{B})$ are the prime dual ideals satisfying condition $I - \sigma = II - \sigma$.*

The proof of Lemma 4.2 is brief. Let $\{a_\iota, \iota \in I\}$, $\bar{I} < \mathfrak{M}$, be a subset of elements of a $\mathfrak{B}(II - \mathfrak{M}, D)$ algebra $\mathfrak{B}$. If $\vee_{\iota\epsilon I}\, a_\iota \neq 1$, there exists element a_0 of $\mathfrak{B}$, $a_0 \neq 0$, with $a_0 < \bar{a}_\iota$ for all ι in I. However, for each non-zero element a_0 of a $\mathfrak{B}(II - \mathfrak{M}, D)$ algebra, there exists a prime dual ideal α_0 of that algebra containing a_0 and in which condition $II - \mathfrak{M}$ is verified. Then $\{\bar{a}_\iota, \iota \in I\} \subseteq \alpha_0$, so that $\wedge_{\iota\epsilon I}\, \bar{a}_\iota$ and thus $\vee_{\iota\epsilon I}\, a_\iota$ exists and the lemma is proved. Referring to statement (A) of the second section, it is now clear that the Boolean algebras isomorphic to $\mathfrak{M}$-fields of sets are exactly the $\mathfrak{B}(II - \mathfrak{M}, D)$ algebras and that, for each such algebra, the associated $\mathfrak{M}$-field of sets may be taken as the field of open-and-closed subsets of the subspace D of $X(\mathfrak{B})$.

Lemma 4.3 is a particular instance of a more general statement (**3**, Theorem 4.2 (3)) and returns us to the subject of P-spaces. From it one sees that for each P-space Y the Boolean algebra $\mathfrak{B}$ of all open-and-closed subsets of Y

is a $\mathfrak{B}(II - \sigma, D)$ algebra with $\beta Y = X(\mathfrak{B})$ and that the space υY may be identified with the subspace D of the representation space of this algebra. However, such $\mathfrak{B}(II - \sigma, D)$ algebras $\mathfrak{B}$ are still of a special character in that $\beta D = X(\mathfrak{B})$. This may be cared for in the following way.

Henceforth, a ***P-Boolean algebra*** will be understood as any $\mathfrak{B}(II - \sigma, D)$ algebra $\mathfrak{B}$ in which the following completeness condition obtains: every collection $\{a_\iota, \iota \in I\}$ of elements of $\mathfrak{B}$ such that each prime dual ideal in D either contains an element of that collection or contains an element of $\mathfrak{B}$ disjoint from every element of the collection has a least upper bound $\bigvee_{\iota \epsilon I} a_\iota$ in $\mathfrak{B}$.

The significance of this completeness condition is explained in two steps. Let $\mathfrak{F}(D)$ denote the field of open-and-closed subsets of the subspace D of the representation space $X(\mathfrak{B})$ of a $\mathfrak{B}(II - \sigma, D)$ algebra $\mathfrak{B}$. As noted in reference to Lemma 4.3, $\mathfrak{F}(D)$ is a reduced, σ-complete field of sets isomorphic to the algebra $\mathfrak{B}$. As the first step, it is shown that $\mathfrak{F}(D)$ is union-intersection closed exactly when the given $\mathfrak{B}(II - \sigma, D)$ algebra satisfies the stated completeness condition. Recall that elements a of $\mathfrak{B}$ are in $1 - 1$ order preserving correspondence with elements O of $\mathfrak{F}(D)$ through the relationship $X(a) \cap D = O$. Then, for any subset $A = \bigcup_{\iota \epsilon I} O_\iota = \bigcap_{j \epsilon J} O_j$ of D, the elements a_ι of $\mathfrak{B}$ corresponding to the elements O_ι in $\bigcup_{\iota \epsilon I} O_\iota$ are such that each prime dual ideal of D in A contains one of the a_ι, while each prime dual ideal of D in $D - A$ contains an element b_j of $\mathfrak{B}$ disjoint from each of the a_ι, namely, an element b_j of $\mathfrak{B}$ corresponding to the complement in D of some O_j in $\bigcap_{j \epsilon J} O_j$. Then, with $a = \bigvee_{\iota \epsilon I} a_\iota$ existing in $\mathfrak{B}$, it is clear that $X(a) \cap D = A$, so that $\mathfrak{F}(D)$ is union-intersection closed. Conversely, if the set $\mathfrak{F}(D)$ is union-intersection closed and $\{a_\iota, \iota \in I\}$ is a family of elements of $\mathfrak{B}$ such that each prime dual ideal in D either contains an element a_ι of this family or an element b_j of $\mathfrak{B}$ disjoint from every member of the family, then, with $A = \bigcup_{\iota \epsilon I}[X(a_\iota) \cap D]$, one has $D - A = \bigcup_{j \epsilon J}[X(b_j) \cap D]$. Then $A = \bigcup_{\iota \epsilon I}[X(a_\iota) \cap D] = \bigcap_{j \epsilon J}[X(\bar{b}_j) \cap D]$. Finally, with a_0 in $\mathfrak{B}$ such that $X(a_0) \cap D = A$, it easily follows that $a_0 = \bigvee_{\iota \epsilon I} a_\iota$ in $\mathfrak{B}$, so that the completeness condition follows.

As the second step, it is now shown that the demand that $\mathfrak{F}(D)$ be union-intersection closed is equivalent to the demand that $\beta D = X(\mathfrak{B})$. First assume that $\mathfrak{F}(D)$ is union-intersection closed. The space $(D, \mathfrak{T})$ consisting of the set D and the topology $\mathfrak{T}$ derived from the field $\mathfrak{F}(D)$ is homeomorphic to the space D as a subspace of $X(\mathfrak{B})$. Hence $\beta(D, \mathfrak{T}) = \beta D$. However, $(D, \mathfrak{T})$ is a P-space so that $\beta(D, \mathfrak{T})$ is the representation space of the algebra of all open-and-closed subsets of $(D, \mathfrak{T})$. With $\mathfrak{F}(D)$ union-intersection closed, this latter algebra is isomorphic to the algebra $\mathfrak{F}(D)$ and thus to the given $\mathfrak{B}(II - \sigma, D)$ algebra $\mathfrak{B}$. Hence $\beta(D, \mathfrak{T}) = X(\mathfrak{B})$. Thus, if $\mathfrak{F}(D)$ is union-intersection closed, then $\beta D = X(\mathfrak{B})$. Conversely, if $\beta D = X(\mathfrak{B})$ so that each open-and-closed subset of D in its relative topology is of the form $X(a) \cap D$, then $\mathfrak{F}(D)$ is obviously union-intersection closed.

The preceding observations are now summarized.

THEOREM 4.4. *The class of all P-Boolean algebras is identical with the class of all algebras of the open-and-closed subsets of the P-spaces. For any P-space Y, the spaces βY and υY are homeomorphic to the spaces $X(\mathfrak{B})$ and D associated with the P-Boolean algebra of all open-and-closed subsets of Y. Two P-spaces Y and Z correspond to the same P-Boolean algebra if, and only if, $\beta Y = \beta Z$.*

With P-spaces characterized as completely regular spaces in which countable intersections of open sets are open, it seems proper to ask concerning completely regular spaces in which any $\mathfrak{M}$-intersection of open sets is open, $\mathfrak{M}$ being a cardinal number presumably larger than $\aleph_0 = \sigma$. Such spaces may be referred to as P-$\mathfrak{M}$-spaces. Let a $\mathfrak{B}(II - \mathfrak{M}, D)$ algebra satisfying the additional completeness condition cited above for P-Boolean algebras be called a P-$\mathfrak{M}$-Boolean algebra. An exact analogue of Theorem 4.4. may then be stated concerning the relationship of P-$\mathfrak{M}$-spaces and P-$\mathfrak{M}$-Boolean algebras.

5. The P'-spaces. The P'-spaces form the second class of completely regular spaces to be discussed here. Their characterization embodied a slight weakening of that of the P-spaces. However, the most enlightening characteristic of the P'-spaces is the following: for each element f of $C(Y)$ and for each point p_0 of $Z(f)$, if there is no neighbourhood U of p_0 in Y such that $f(p) = 0$ throughout U, then there is a deleted neighbourhood U' of p_0 such that $f(p) > 0$ throughout U' or $f(p) < 0$ throughout U'. It is this feature of P'-spaces that guides the next procedures. Use is also made of the fact (**4**, Theorem 8.4) that, for each P'-space Y, $\beta Y = X(\mathfrak{B})$ where $\mathfrak{B}$ is a σ-complete Boolean algebra.

Let a point p_0 of an arbitrary completely regular space Y be termed a P'-point of Y if it has the property cited just above.

LEMMA 5.1. *Let Y be a completely regular space such that $\beta Y = X(\mathfrak{B})$ where $\mathfrak{B}$ is a σ-complete Boolean algebra. Let each point p of Y as in $X(\mathfrak{B})$ be considered as a prime dual ideal α_p of $\mathfrak{B}$. Then a point $\bar{p}$ of Y is a P'-point of Y if, and only if, the corresponding prime dual ideal $\alpha_{\bar{p}}$ satisfies the following condition: for each countable union $1 = \vee a_n$ in $\mathfrak{B}$ of which no component element a_n is in $\alpha_{\bar{p}}$, there exists a non-zero element a_0 of $\mathfrak{B}$ with a_0 in $\alpha_{\bar{p}}$ and such that all other α_p containing a_0 contain likewise some component of the given union.*

Proof. Assume first that $\bar{p}$ is a P'-point of Y. Let $1 = \vee a_n$ be a disjoint countable union of elements of $\mathfrak{B}$ of which no component a_n is in $\alpha_{\bar{p}}$. Then, because of the σ-completeness of $\mathfrak{B}$, there exists an element f of $C(X[\mathfrak{B}])$ with $f(\alpha) = 1/n$ for each prime dual ideal (point) α containing a_n. Now let a_0 be any element of $\mathfrak{B}$ in $\alpha_{\bar{p}}$. Then $a_0 \wedge a_n \neq 0$ for at least one element a_n of the union $1 = \vee a_n$ and, since α_p contains no element of this union, actually $a_0 \wedge a_n \neq 0$ for infinitely many subscripts n. From this it follows that $f(\alpha_{\bar{p}}) = 0$. However, with $\bar{p}$ a P'-point of Y, there exists a deleted neighbourhood U'

of $\bar{p}$ in Y and thus a particular element a_0 of $\mathfrak{B}$ in $\alpha_{\bar{p}}$ such that $f(\alpha_p) > 0$ for all α_p containing a_0, $\alpha_p \neq \alpha_{\bar{p}}$. However, $f(\alpha_p) > 0$ means $f(\alpha_p) = 1/n$ for some n. This, in turn, is easily seen to mean that $a_n \in \alpha_p$. Thus there exists an element a_0 of $\mathfrak{B}$ in $\alpha_{\bar{p}}$ such that every α_p containing a_0, $\alpha_p \neq \alpha_{\bar{p}}$, contains likewise some element a_n of the given countable union.

Conversely, assume that prime dual ideal $\alpha_{\bar{p}}$ of σ-complete algebra $\mathfrak{B}$ corresponding to point $\bar{p}$ of Y has the property with respect to countable unions stated in the theorem. Let element f of $C(Y)$ be such that $f(\bar{p}) = 0$. Assume, for the moment, that f is non-negative throughout Y. Let $f_0 = f \wedge 1$ in the usual sense of function lattices. Let $\bar{f}_0$ or, for notational simplicity, simply f denote the extension of f_0 over $\beta Y = X(\mathfrak{B})$. Let $O_n = [\alpha \in X(\mathfrak{B}) \mid f(\alpha) < 1/n]$. Then, by reason of the σ-completeness of $\mathfrak{B}$, there exists element a_n of $\mathfrak{B}$ such that $X(a_n) = \bar{O}_n$. The sequence $\{a_n\}$ is obviously such that $a_{n+1} < a_n$. Form the element $a_0 = \wedge a_n$ in $\mathfrak{B}$. Finally, construct a new sequence $\{b_n\}$ in $\mathfrak{B}$ with: $b_0 = \bar{a}_0$, $b_1 = 1 \wedge \bar{a}_1$, $b_2 = a_1 \wedge \bar{a}_2, \ldots$.

Now $\vee_{n=0}^{\infty} b_n = 1$ and is a countable disjoint union. If some (non-zero) b_n is in $\alpha_{\bar{p}}$, clearly this b_n is $b_0 = a_0$ and one concludes that $f(\alpha_p) = 0$ for all α_p with $a_0 \in \alpha_p$. Then $U = [p \in Y \mid a_0 \in \alpha_p]$ is a neighbourhood of $\bar{p}$ in Y such that $f(p) = 0$ throughout U. If no b_n is in $\alpha_{\bar{p}}$, then, by hypothesis, there is an element c_0 of $\mathfrak{B}$ in $\alpha_{\bar{p}}$ such that every α_p containing c_0, $\alpha_p \neq \alpha_{\bar{p}}$, contains some (non-zero) b_n. Since b_0 is here assumed as not contained in $\alpha_{\bar{p}}$, this first c_0 may be replaced by $\bar{b}_0 \wedge c_0$. Denote this element also by the symbol c_0. Then each α_p containing c_0, $\alpha_p \neq \alpha_{\bar{p}}$, contains also an element b_n of the countable disjoint union and this b_n is not the element b_0. However, with $b_n = a_{n-1} \wedge \bar{a}_n$ in α_p, $n > 1$, then $1/n < f(\alpha_p) < 1/(n-1)$ so that $f(\alpha_p)$ is non-zero. One concludes from this that $U' = [p \in Y \mid p \neq \bar{p}$ and $c_0 \in \alpha_p]$ is a deleted neighbourhood of $\bar{p}$ in Y such that $f(p) > 0$ throughout U'.

Finally, for an arbitrary element f of $C(Y)$ with $f(\bar{p}) = 0$, first apply the above analysis to the elements f^+, f^- formed in the usual function-lattice sense. Note that if $f^+(p) > 0$ throughout a deleted neighbourhood, then $f^-(p) = 0$ throughout the same neighbourhood. With this in mind, this converse part of the theorem is easily seen to hold for all elements f of $C(Y)$ with $f(\bar{p}) = 0$.

THEOREM 5.2. *Let $X(\mathfrak{B})$ be the Stone representation space of a σ-complete Boolean algebra $\mathfrak{B}$. Let Y be a subspace of $X(\mathfrak{B})$ such that $\beta Y = X(\mathfrak{B})$ and also such that for every countable union $\vee a_n = 1$ in $\mathfrak{B}$ each point (prime dual ideal) α_0 of Y either contains a component of this union or contains an element a_0 of $\mathfrak{B}$ such that every other point α of Y which contains a_0 contains an element of this union. Then Y is a P'-space and every P'-space may be thus described.*

For the sake of brevity, a Boolean algebra of the type described in Theorem 5.2 will be called a P'-Boolean algebra. The description of such algebras is very awkward. However, with $\mathfrak{B}$, $X(\mathfrak{B})$ and Y as described in that theorem, consider the field $\mathfrak{F}(Y)$ of open-and-closed subsets of Y. Obviously $\mathfrak{F}(Y)$ is

reduced and union-intersection closed. In view of the σ-completeness of $\mathfrak{B}$, also $\mathfrak{F}(Y)$ is σ-complete in the sense that every countable set of elements of $\mathfrak{F}(Y)$ is contained in a smallest element of $\mathfrak{F}(Y)$. Finally, from Theorem 5.2, $\mathfrak{F}(Y)$ is seen to have an additional property that may be called the near-σ-field property; if O is the smallest element of $\mathfrak{F}(Y)$ including each of the elements $\{O_n\}$ and if point $\bar{p}$ of Y is in O but in no O_n, then there exists element O_0 of $\mathfrak{F}(Y)$ with $\bar{p} \in O_0$ while $p \in O_0$, $p \neq \bar{p}$, implies $p \in O_n$ for some n. Thus for any P'-Boolean algebra $\mathfrak{B}$ as described in Theorem 5.2 the associated field $\mathfrak{F}(Y)$ is a reduced, union-intersection closed, σ-complete, near-σ-field of sets which, as a Boolean algebra, is isomorphic to $\mathfrak{B}$ while the space $(Y, \mathfrak{T})$ derived from $\mathfrak{F}(Y)$ is homeomorphic to the P'-space Y. Note that $\beta(Y, \mathfrak{T})$, as homeomorphic to $X(\mathfrak{B})$, is of dimension zero.

Conversely, let $\mathfrak{F}(Y)$ be a reduced, union-intersection closed, σ-complete, near-σ-field of sets and let $(Y, \mathfrak{T})$ be formed as usual. Then, by methods similar to those used in Theorem 3.1, it may be proved that $(Y, \mathfrak{T})$ is a P'-space, provided one has assurance that $\beta(Y,\mathfrak{T})$ is of dimension zero. Whether or not such assurance is contained in the stated assumptions regarding $\mathfrak{F}(Y)$, the present writer does not know. However, he has indicated elsewhere **(6)** how to state such assurance regarding $\beta(Y, \mathfrak{T})$ in purely set-theoretic language.

These observations are now summarized.

THEOREM 5.3. *The P'-Boolean algebras are identical with the algebras formed under the inclusion relation by elements of reduced, union-intersection closed, σ-complete, near-σ-fields of sets $\mathfrak{F}(Y)$ with $\beta(Y, \mathfrak{T})$ of dimension zero. Such fields, in turn, may be identified with the fields of open-and-closed subsets of the P'-spaces.*

6. The *UF*-Boolean algebras. We turn now to the U-spaces described in **(4)**. A completely regular space X is a U-space if, and only if, to each element f of $C(X)$ there is associated a unit element u in $C(X)$ such that $f = u \,.\, |f|$. For any completely regular space X, X is a U-space if, and only if, βX is a U-space **(4**, Theorem 5.2**)**. Finally, βX is a U-space if, and only if, it is zero-dimensional and for each element f of $C(\beta X)$ the sets $P(f)$ and $N(f)$ are completely separated in βX. The zero-dimensionality of such βX links the U-spaces to Boolean algebras.

Let $\mathfrak{B}$ again denote an arbitrary Boolean algebra. Let $\rho = \{a_n\}$ denote a monotone, non-decreasing sequence of elements of $\mathfrak{B}$. For the sake of brevity, refer to a sequence like ρ as a *tower* in $\mathfrak{B}$. Two towers $\rho = \{a_n\}$ and $\tau = \{b_n\}$ will be called disjoint if $a_n \wedge b_n = 0$ for each positive integer n. Finally, an element a_0 of $\mathfrak{B}$ will be called a *cap* of a tower ρ if $a_n < a_0$ for each element a_n of $\rho = \{a_n\}$.

Now define a Boolean algebra $\mathfrak{B}$ to be a *UF-Boolean algebra* if, and only if, disjoint towers in $\mathfrak{B}$ have disjoint caps in $\mathfrak{B}$. The *UF*-Boolean algebras have a close relationship to the U-spaces (and F-spaces) of **(3**; **4)**.

THEOREM 6.1. *The UF-Boolean algebras are exactly those Boolean algebras $\mathfrak{B}$ for which the sets $P(f)$ and $N(f)$ are completely separated in $X(\mathfrak{B})$ for each element f of $C[X(\mathfrak{B})]$.*

Proof. Assume that $\mathfrak{B}$ is a *UF*-Boolean algebra and let f be an element of $C[X(\mathfrak{B})]$. Let $F_n = [\alpha \in X(\mathfrak{B}) \mid f(\alpha) \geqslant 1/n]$ while $O_n = [\alpha \in X(\mathfrak{B}) \mid f(\alpha) > 1/(n + \frac{1}{2})]$. Then, using the compactness of F_n and the openness of O_n, one can conclude to the existence in $\mathfrak{B}$ of an element a_n such that $F_n \subseteq [\alpha \in X(\mathfrak{B}) \mid a_n \in \alpha] \subseteq O_n$. Moreover, since $F_n \subseteq O_n \subseteq F_{n+1} \subseteq O_{n+1}$, one has $a_n \leqslant a_{n+1}$ and the sequence $\rho = \{a_n\}$ is a tower in $\mathfrak{B}$. Similarly, with $F_n^* = [\alpha \in X(\mathfrak{B}) \mid f(\alpha) \leqslant -1/n]$ and $O_n^* = [\alpha \in X(\mathfrak{B}) \mid f(\alpha) < -1/(n + \frac{1}{2})]$, let a second tower $\tau = \{b_n\}$ be constructed with $F_n^* \subseteq [\alpha \in X(\mathfrak{B}) \mid b_n \in \alpha] \subseteq O_n^*$. The two towers thus formed are clearly disjoint and thus, by assumption, have disjoint caps a_0 and b_0. It is now but a small matter to verify that $P(f) \subseteq [\alpha \in X(\mathfrak{B}) \mid a_0 \in \alpha]$ and $N(f) \subseteq [\alpha \in X(\mathfrak{B}) \mid b_0 \in \alpha]$ so that the sets $P(f)$ and $N(f)$ are completely separated in $X(\mathfrak{B})$.

Conversely, assume that for each element f of $C[X(\mathfrak{B})]$ the sets $P(f)$ and $N(f)$ are completely separated in $X(\mathfrak{B})$. Let $\rho = \{a_n\}$ and $\tau = \{b_n\}$ be a pair of disjoint towers in $\mathfrak{B}$. Let f_n be the unique element of $C[X(\mathfrak{B})]$ with $f_n(\alpha) = 1$ for all α with $a_n \in \alpha$, with $f_n(\alpha) = -1$ for all α with $b_n \in \alpha$ and with $f_n(\alpha) = 0$ for all α containing $\bar{a}_n \wedge \bar{b}_n$. Finally, form $f_0 = \Sigma_{n=1}^{\infty} f_n/2^n$. Then f_0 is an element of $C[X(\mathfrak{B})]$ and, by assumption, the sets $P(f_0)$ and $N(f_0)$ are completely separated in $X(\mathfrak{B})$. In virtue of the zero-dimensionality of $X(\mathfrak{B})$, this implies that there exists elements a_0 of $\mathfrak{B}$ such that $P(f_0) \subseteq [\alpha \in X(\mathfrak{B}) \mid a_0 \in \alpha]$, while $N(f_0) \subseteq [\alpha \in X(\mathfrak{B}) \mid \bar{a}_0 \in \alpha]$. The element a_0 is now seen to cap the tower $\rho = \{a_n\}$ while its complement $\bar{a}_0$ caps the tower $\tau = \{b_n\}$. Thus the theorem is proved.

The observations of this section may now be summarized.

THEOREM 6.2. *Any UF-Boolean algebra is the algebra of all open-and-closed subsets of some U-space and any such algebra is a UF-Boolean algebra. Two U-spaces Y and Z correspond to the same UF-Boolean algebra if, and only if, $\beta Y = \beta Z$.*

7. Comments. This section begins with an observation concerning *F*-spaces (**4**). A completely regular space Y is an *F*-space if, and only if, for each element f of $C(Y)$ the sets $\bar{P}(f)$ and $\bar{N}(f)$ are completely separated. Every *F*-space Y has the following property (**4**, Theorem 2.6) pertinent to our purpose: for each zero set Z of Y each element f of $C^*(Y - Z)$ has a continuous extension $\bar{f}$ in $C^*(Y)$. Here $C^*(Y)$ indicates the collection of bounded elements of $C(Y)$.

LEMMA 7.1. *Let Y be a completely regular F-space. Then βY is without G_δ-points other than isolated points. Moreover, a point p of Y is a non-isolated*

G_δ-point in Y if, and only if, every element f of $C^*(Y - \{p\})$ has a continuous extension at p while some element of $C(Y - \{p\})$ lacks such an extension.

Proof. As regards the first assertion, assume that p is a G_δ-point of βY. If p is not an imbedded point of Y in βY, then every element of $C^*(\beta Y - \{p\})$ has a continuous extension at p by definition of βY. If p is an imbedded point of Y in βY, then $\{p\}$ is a zero set in Y and, by the property of F-spaces cited above, one again concludes that every element of $C^*(\beta Y - \{p\})$ has a continuous extension at p. Hence $\beta(\beta Y - \{p\}) = \beta Y$ unless p is an isolated point of βY. However, for any completely regular space X the cardinality of a zero set contained in $\beta X - X$ is at least exp (exp $\aleph_0$) (**7**, Theorem 49). Thus the point p must be an isolated point in βY.

As to the second assertion, it is merely to be noted that if a point p of Y has the extension properties listed in the theorem, then $\beta(Y - \{p\}) = \beta Y$ while $p \notin v(Y - \{p\})$. From this it follows easily that such a point is a G_δ-point (**5**, Example 2.3).

Now the F-spaces X such that βX is zero-dimensional and thus of present interest are identical with the U-spaces (**4**, Theorem 5.5). With the U-spaces described in terms of Boolean algebras, attention may now be called to the following conclusion.

THEOREM 7.2. *The Stone representation spaces of Boolean σ-algebras and, more generally, of UF-Boolean algebras are without G_δ-points other than isolated points.*

This theorem cannot be extended to include all Boolean algebras. In a written communication, C. W. Kohls called the attention of the writer to the following example.

Example. Let N denote the set of all positive integers. Let $\mathfrak{B}(N)$ denote the class of all finite subsets of N along with their complements in N together with the empty set and the set N itself. As partially ordered by the inclusion relation, $\mathfrak{B}(N)$ is a Boolean algebra. In $X(\mathfrak{B}[N])$ there is only one prime dual ideal other than the point-principal dual ideals. That ideal consists of all the infinite subsets of N in $\mathfrak{B}(N)$. As a point of $X(\mathfrak{B}[N])$ this ideal is obviously a non-isolated G_δ-point. It is also easily seen that $\mathfrak{B}(N)$ is not a UF-Boolean algebra. Thus let $a_n = \{1, 3, \ldots, 2n - 1\}$ and $b_n = \{2, 4, \ldots, 2n\}$. Then, as elements of $\mathfrak{B}(N)$, $a_n < a_{n+1}$, $b_n < b_{n+1}$ and $a_n \wedge b_n = 0$. However, it is impossible to find in $\mathfrak{B}(N)$ elements a_0, b_0 with $a_0 \wedge b_0 = 0$ and such that $a_n < a_0$ and $b_n < b_0$ for all positive integers n.

References

1. C. C. Chang, *On the representation of α-complete Boolean algebras*, Trans. Amer. Math. Soc., *85* (1957), 208–18.
2. S. Enomoto, *Boolean algebras and fields of sets*, Osaka Math. J., *5* (1953), 99–115.
3. L. Gillman and M. Henriksen, *Concerning rings of continuous functions*, Trans. Amer. Math. Soc., 77 (1954), 340–62.
4. ——— *Rings of continuous functions in which every finitely generated ideal is principal*, Trans. Amer. Math. Soc., *82* (1956), 362–91.
5. L. J. Heider, *A note concerning completely regular G_δ spaces*, Proc. Amer. Math. Soc., *8* (1957), 1060–6.
6. ——— *Compactifications of dimension zero*, Notices Amer. Math. Soc., 552–9.
7. E. Hewitt, *Rings of real-valued continuous functions, I*, Trans. Amer. Math. Soc., *64* (1948), 45–99.
8. J. R. Isbell, *Zero-dimensional spaces*, Tohoku Math. J., *7* (1955), 1–8.
9. R. S. Pierce, *Distributivity in Boolean algebras*, Pacific J. Math., 7 (1957), 983–93.
10. R. Sikorski, *On the representation of Boolean algebras as fields of sets*, Fund. Math., *35* (1948), 247–58.
11. E. C. Smith, Jr., *A distributivity condition for Boolean algebras*, Ann. Math., *64* (1956), 551–61.
12. M. H. Stone, *Boundedness properties in function lattices*, Can. J. Math., *1* (1949), 176–86.

Institute for Advanced Study
Marquette University

ON A PAPER OF MAURICE SION

MARK MAHOWALD

1. Let M_0 be the set of measures μ on the real line such that open sets are μ^*-measurable. While attempting to find out whether a set μ^*-measurable for all μ in M_0 is mapped into a similar set by a continuous function of bounded variation, Maurice Sion develops a theory for what he calls variational measure **(4)**. As an application of the theory, he gets conditions on a function f and a set of measures M in order that f map a set, which is μ^*-measurable for all $\mu \in M$, into a set of the same kind. In particular he proves for his class M_2 (def. 2.5), the following theorem **(4**, § 8.11**)**.

THEOREM. *If A is measurable for all measures in M_2 and if f is continuous from the irrationals to* $[0, 1]$, *then $f(A)$ is measurable for all measures in M_2.*

Since all projective sets are continuous images of the irrationals **(2**, p. 39**)** and since the existence of a non-measurable projective set is consistent with the axioms of set theory if they are consistent, **(1)**, Sion concludes that Lebesgue measure is not in M_2.

We prove Sion's result in another way and more importantly, we characterize M_2 completely with respect to open regular measures. As an application, we prove, without the continuum hypothesis, the existence of a function discontinuous on every set of positive outer measure (Lebesgue).

The author is indebted to William Larkin for many helpful critical comments.

2. Notation and definitions.

2.1. A *partition*, $P(S)$, of a set S is a collection of sets, $E \subset S$, finite in number, pairwise disjoint and whose union is S.

2.2. A *refinement of a partition*, P_1, is a second partition, P_2, such that each set in P_2 is a subset of some set in P_1.

2.3. An *open regular measure* is a measure such that each μ^*-measurable set has a measurable cover which is a G_δ set.

2.4. $M_0 = |\mu\colon \mu$ is a measure on $[0, 1]$ and open sets are μ^*-measurable$|$.

2.5. A sequence with property A is a sequence of partitions $P_n(S)$ such that

(a) $S \subset [0, 1]$ and $0 < \mu^*(S) < \infty$;
(b) P_{n+1} is a refinement of P_n;
(c) if $B \subset S$ and $\mu^*(B) > 0$, then

Received October 17, 1958.

$$\lim_{n\to\infty} \sum_{E \in P_n} \mu^*(B \cap E) = \infty.$$

2.6. $M_2 = \{\mu : \mu \in M_0$ and there does not exist a set S with a partition sequence having property $A\}$.

2.7. $M_3 = \{\mu : \mu \in M_0$, μ is open regular, there exists a partition sequence $P_n([0, 1])$ with property $A\}$.

2.8. A measure will be called *non atomic* if no single point has positive outer measure.

3. Conditions implying a measure is not in M_2 or is in M_3.

3.1. THEOREM. *If there exists a set $S \subset [0, 1]$ of positive outer measure and a bounded function f defined on S discontinuous on every set $E \subset S$ for which $\mu^*(E) > 0$ and if μ is open regular, then $\mu \notin M_2$.*

The proof is long and will be given in § 5.

3.2. COROLLARY. *If $S = [0, 1]$ in the theorem, then $\mu \in M_3$.*

This corollary is an immediate consequence of the proof of the theorem (see 5).

3.3. The following lemma is obtained by a minor modification of the proof of the similar theorem (without the word "bounded") due to Sierpinski and Zygmund **(3)**.

LEMMA. *There exists a bounded function from the reals to the reals which is discontinuous on every set having the power of the continuum.*

3.4. Then we can prove this

THEOREM. *If μ is such that every set of positive outer measure has the power of the continuum and μ is open regular, then μ is in M_3.*

Proof. The theorem follows immediately from 3.2 and 3.3.

3.5. COROLLARY. *If there exists one set of positive outer measure such that all subsets of positive outer measure have the power of the continuum and if μ is open regular for all subsets of this set, then $\mu \notin M_2$.*

3.6. COROLLARY. *Under the continuum hypothesis: If μ is non-atomic and open regular, then μ is in M_3. If there exists a subset of positive outer measure such that every single point subset has measure zero, then $\mu \notin M_2$.*

Proof. If a measure is non-atomic then every countable set has measure zero. The continuum hypothesis then implies every set having positive outer measure has the power of the continuum and 3.4 and 3.5 prove the theorem.

3.7. Every measure on the subsets of the unit interval is either in M_3 or it is not. The definition of M_3, which enables one to decide whether or not a measure is in M_3, does not depend on the continuum hypothesis, that is, the

definition makes sense if the hypothesis is true or false. Now, if there exists a non-atomic, open regular measure not in M_3, then this can be shown by a set theoretic argument. Such an argument with corollary 3.6 would be a proof from set theory of the proposition: *the continuum hypothesis is false.* Gödel **(1)** has shown that this cannot be proven with such an argument. Therefore, all open regular, non-atomic measures are in M_3, that is, we can improve 3.6 to the following

THEOREM. *If μ is open regular and non-atomic, then $\mu \in M_3$.*

3.8. We can restate this by this

THEOREM. *There are no σ-finite open regular measures in M_2. Lebesgue measure is not in M_2 but it is in M_3.*

4. A converse to Theorem 3.1.

4.1. THEOREM. *If μ is an open regular measure not in M_2 and S is a set with a partition sequence having property A, then there is a function defined on S which is discontinuous on every subset E of S such that $\mu^*(E) > 0$.*

Proof. For μ, there exists a sequence $P_m(S)$ of partitions with property A. Let $F_{11}, \ldots, F_{k1}$ be a numbering of the sets of P_1. Let n_1 be the smallest integer larger than $\log_2 k$. Define

$$f_1(x) = (i-1)/2^{n_1}$$

for $x \in F_{i1}$, for $i = 1, \ldots, k$.

Suppose for $m-1$ we have defined n_{m-1}, a numbering, $F_{i,m-1}$, for the partition P_{m-1}, and f_{m-1}. The induction step will be defined as follows:
Let $p_m = \max_q j_q$, where j_q is the number of sets in P_m which are subsets of $F_{q,m-1}$. Let h_m be the smallest integer greater than $\log_2(p_m + 2)$. Let $n_m = h_m + n_{m-1}$. Let F_{im}, for

$$i = (q-1)2^{h_m} + 1, \ldots, (q-1)2^{h_m} + j_q$$

and

$$q = 1, \ldots, 2^{n_{m-1}},$$

be a numbering of the sets of P_m which are subsets of $F_{q,m-1}$. If F_{im} does not appear in this numbering, then $F_{im} = \phi$. Then define

$$f_m(x) = (i-1)/2^{n_m} \qquad \text{for} \qquad x \in F_{im}.$$

The sequence f_m is monotonically non-decreasing and is uniformly bounded by one. Therefore there exists a limit function f_0.

For m fixed, our choice of h_m assures us that

$$|x : f_m(x) = (q2^{h_m} - 1)/2^{n_m}| = \phi$$

since it would equal

$$F_{q2^{h_m},m}$$

which is empty. As a consequence, we have: if

$$f_0(x) < (q2^{h_m})/2^{n_m}$$

for any m and

$$q = 1, \ldots, 2^{n_m - 1},$$

then

$$f_0(x) \leq (q2^{h_m} - 1)/2^{n_m}.$$

Therefore

$$F_{im} = |x: [i - 1 - (1/2^{n_m+1})]/2^{n_m} < f_0(x) < i/2^{n_m}|.$$

Now suppose there exists a set $B \subset S$ such that $\mu^*(B) > 0$ and such that f_0 is continuous on B. Since F_{im} is the inverse image of an open set, $F_{im} \cap B$ is open in B, that is, there exists an open set U_{im} such that $F_{im} \cap B = U_{im} \cap B$. Let $U_{ijm} = U_{im} \cap U_{jm}$ for $i \neq j$ and $U_{jjm} = \phi$ and let $V_{im} = U_{im} - \bigcup_j U_{ijm}$. Clearly V is pairwise disjoint for m fixed. Also, since F is pairwise disjoint for m fixed, no point of $F_{jm} \cap B$ can be in U_{im} for $i \neq j$. Therefore, we have $V_{im} \cap B = F_{im} \cap B$. Hence we can choose a measurable cover of $F_{im} \cap B$, C_{im}, which is a subset of V_{im}. Therefore, C is pairwise disjoint and

$$\sum_i \mu^*(F_{im} \cap B) = \sum_i \mu(C_{im}) = \mu(U_i C_{im}) = \mu^*(B).$$

Since m is arbitrary, we have

$$\lim_{m \to \infty} \sum_i \mu^*(F_{im} \cap B) = \mu^*(B)$$

but this is just

$$\lim_{m \to \infty} \sum_{E \in P_m} \mu^*(B \cap E) = \mu^*(B) \leq \mu^*(S) < \infty.$$

This contradiction proves the theorem:

4.2. COROLLARY. *For every $\mu \in M_3$, there exists a function discontinuous on every set having positive outer measure. In particular there exists such a function for Lebesgue measure.*

5. Proof of Theorem 3.1. Let f be the function described in the theorem. Since f is bounded we can suppose that $0 \leq f(x) \leq 1$. Let $E_{ni} = |x: i/2^n \leq f(x) < (i+1)/2^n|$, $i = 0, \ldots, 2^n$. Set $P_n(S) = \{E_{ni}\}$; we shall prove that this sequence has the property A. The facts that P is a partition and that P_{n+1} is a refinement of P_n are clear. We need only show that, for any $B \subset S$ for which $\mu^*(B) > 0$, we have

$$\lim_{n \to \infty} \sum_{i=0}^{2^n} \mu^*(E_{ni} \cap E) = \infty.$$

Assume that there exists a set $E \subset S$ such that

$$0 < \lim_{n \to \infty} \sum_i \mu^*(E_{ni} \cap E) = a < \infty. \tag{1}$$

We then shall prove that there exists a subset of E having positive outer measure and on which f is continuous and this contradiction will prove the theorem.

Subadditivity of μ^* implies that the limit in (1) approaches a from below. Therefore, there exists an N such that $n > N$ implies, for $a/10 > \epsilon > 0$,

$$a - \epsilon < \sum_i \mu^*(E_{ni} \cap E) < a$$

and in particular

$$a - \epsilon < \sum_j \mu^*(E_{Nj} \cap E) < a.$$

Let B_{Nj} be a measurable cover of $E \cap E_{Nj}$. If E_{ni} is a subset of E_{Nj}, we shall write E_{nij} and we shall designate a measurable cover of $E \cap E_{nij}$ by B_{nij}. It is easily shown that the sets B_{nij}, $n > N$, can be so determined that if I is a set of integers such that $\bigcup_{i \in I} E_{mi} = E_{nk}$, then $\bigcup_{i \in I} B_{mij} = B_{nkj}$.

We next derive measurable sets H_{nij} contained in B_{nij}, disjoint for each fixed pair n, j and with

$$a - 2\epsilon < \sum_{j,i} H_{nij} < a.$$

Let

$$B_{nikj} = B_{nij} \cap B_{nkj} \qquad i \neq k \tag{2}$$
$$\phantom{B_{nikj}} = \phi \qquad i = k$$

and let $H_{nij} = B_{nij} - \bigcup_k B_{nikj}$. Since, for $n > N$,

$$a - \epsilon < \sum_j \mu(B_{Nj}) = \sum_j \mu(\bigcup_i B_{nij}) < a$$

and

$$a - \epsilon < \sum_{j,i} \mu(B_{nij}) < a,$$

we have

$$\sum_j \left[\sum_i \mu(B_{nij}) - \mu(\bigcup_i B_{nij}) \right] \geq \sum_{j,i} \mu(\bigcup_k B_{nikj}) < \epsilon.$$

From this and the definition of H, we have

$$a - 2\epsilon < \sum_{j,i} \mu(H_{nij}) < a \tag{3}$$

for all $n > N$. By the choice of B, $\bigcup_i H_{nij}$ is monotonically decreasing as a function of n for each j. Letting $H_j = \bigcap_n \bigcup_i H_{nij}$, we have from (3)

$$\sum_j \mu(H_j) > a - 2\epsilon. \tag{4}$$

We next obtain formulas analagous to (3) and (4) with the sets H_{nij} replaced by open sets. Let V_{nij} be an open cover of H_{nij} such that

$$\mu(V_{nij}) \subset \mu(H_{nij}) + \epsilon/2^i(2^N + 1)$$

and if I is a set of integers such that $\bigcup_{i \in I} E_{mi} = E_{nk}$, then $\bigcup_{i \in I} V_{mij} \subset V_{nkj}$. Such a cover exists because of the open regularity of μ. Then for every $n > N$ we have

$$a - 2\epsilon < \sum_{i,j} \mu(H_{nij}) = \sum_j \mu(\bigcup_i H_{nij}) < \sum_j \mu(\bigcup_i V_{nij})$$
$$< \sum_{i,j} \mu(V_{nij}) < \sum_{i,j} \mu(H_{nij}) + \epsilon < a + \epsilon.$$

Therefore

$$\sum_{i,j} \mu(V_{nij}) - \sum_j \mu(\bigcup_i V_{nij}) < 3\epsilon.$$

Using notation analogous to that of (2), letting $U_{nij} = V_{nij} - \bigcup_k V_{nikj}$, and using the same argument which leads to (3) and (4), we have

$$\text{(5)} \qquad \sum_{j,i} \mu(\bigcup_k V_{nikj}) < 3\epsilon \quad \text{and} \quad a - 5\epsilon < \sum_{j,i} \mu(U_{nij}) < a + \epsilon$$

for all $n > N$.

By the choice of V, $\bigcup_i U_{nij}$ is monotonically decreasing as a function of n for each j. Letting $U_j = \bigcap_n \bigcup_i U_{nij}$, we have from (5)

$$\text{(6)} \qquad \lim_{n\to\infty} \sum_{j,i} \mu(U_{nij}) = \sum_j \lim_{n\to\infty} \mu(\bigcup_i U_{nij}) = \sum_j \mu(U_j) > a - 5\epsilon.$$

We shall now show

$$\text{(7)} \qquad \sum_j \mu(U_j \cap H_j) > a - 8\epsilon.$$

Since

$$\mu(V_{Nj}) = \mu(V_{Nj} - (U_j \cup H_j)) + \mu(\mathrm{H}_j \cup U_j)$$

and

$$\mu(H_j \cup U_j) = \mu(\mathrm{H}_j) + \mu(U_j) - \mu(H_j \cap U_j),$$

we have

$$\sum_j [\mu(H_j) + \mu(U_j) - \mu(H_j \cap U_j)] < \sum \mu(V_{Nj}) < a + \epsilon$$

or

$$a - 2\epsilon + a - 5\epsilon - \sum_j \mu(H_j \cap U_j) < a + \epsilon.$$

This yields (7).

Pick a j such that $\mu(U_j \cap H_j) > 0$. Then

$$\mu^*(\mathrm{E} \cap E_{Nj} \cap U_j) > \mu(U_j \cap H_j) > 0.$$

Let $C = E \cap E_{Nj} \cap U_j$. Then for arbitrary but fixed $n > N$ we have

$$V_{nij} \cap V_{nkj} \cap C = \phi \textit{ for } i = k.$$

We shall show that f is continuous on C. Let $\delta > 0$ be given. Then there exists an n such that $2^{-n} < \delta$. Let x_0 be in C. Then

$$|f(x_0) - f(x)| < 2^{-n} < \delta$$

for all x in $V_{nij} \cap C$ where i is such that x_0 is in E_{ni}. Therefore f is continuous on C contrary to hypothesis on f. This contradiction proves the theorem.

References

1. K. Gödel, *The consistency of the axiom of choice and of the generalized continuum-hypothesis*, Proc. Nat. Acad. Sci. U.S.A., *24* (1938), 556–7.
2. W. Sierpinski, *Les ensembles projectifs et analytiques*, Mémorial des Sciences Mathématiques, no. 112 (1950).
3. W. Sierpinski and A. Zygmund, *Sur une fonction qui est discontinue sur tout ensemble de puissance du continu*, Fund. Math., *4* (1923), 316–18.
4. Maurice Sion, *Variational Measure*, Trans. Amer. Math. Soc., *83* (1956), 205–21.

Xavier University

ON GENERALIZED MORSE-TRANSUE FUNCTION SPACES

H. W. Ellis

1. Introduction. Marston Morse and William Transue **(6, 8)** have introduced and studied function spaces, called *MT*-spaces, for which the elements of the topological dual are of integral type. Their theory does not admit certain classical Banach function spaces including spaces of bounded functions and $\mathfrak{L}_C^\infty$ spaces. The theory of function spaces determined by a length function (λ-spaces) **(4, 5)**, which depends on a fixed measure, admits many of the maximal *MT*-spaces, the spaces $\mathfrak{L}_C^\infty$ and spaces of locally integrable functions but does not admit certain maximal *MT*-spaces including the space $\mathfrak{K}_C$ of complex continuous functions with compact supports.

In **(4)** the definition of *MT*-spaces was weakened by dropping the requirement that $\mathfrak{K}_C$ be dense in the space and making no hypothesis concerning the dual. The resulting spaces were called MT^*-spaces and the elements of intregal type in the dual then constituted the *MT*-conjugate of the space. A λ-space **(4)** is an MT^*-space if it contains $\mathfrak{K}_C$. The *MT*-spaces are just those MT^*-spaces for which the dual and *MT*-conjugate coincide. The space of bounded functions on a suitable space E is an MT^*-space that is neither an *MT*- nor a λ-space.

In the development of the theory of *MT*-spaces an important role was played by the fact that the semi-norm $\mathfrak{N}^A$ could be defined in A and extended to all of $\mathbf{C}^E$ by (3.2) below. Since there are MT^*-spaces for which the *MT*-conjugate reduces to the zero element of the dual (§ 3), (3.2) is not valid for every MT^*-space. For an $\mathfrak{N}^A$-extensible MT^*-space (Definition 3.2) (3.2) holds. Since $\mathfrak{N}^A$ is then a reflexive semi-norm, the *MT*-conjugate is then dense in the dual of A in the $\sigma(A', A)$ topology (Theorem 3.1). The $\mathfrak{N}^A$-extensible MT^*-spaces have many of the properties of general *MT*-spaces.

The last part of this paper is mainly concerned with the role played in the general theory of MT^*-spaces by the λ-spaces. When E is countable at infinity this can be simply stated as follows. If A is a λ-space containing $\mathfrak{K}_C$, A is an $\mathfrak{N}^A$-extensible MT^*-space for which every measure in $\mathfrak{A}^*$ is of base μ (Theorem 3.3.). Conversely if A is an $\mathfrak{N}^A$-extensible MT^*-space for which every measure in $\mathfrak{A}^*$ is of base μ, $\mathfrak{N}^A$ extended by (3.2) determines a length function λ (Theorem 4.1) and $\mathfrak{L}_C^\lambda$, the λ-space determined by λ, and Ω^A (§ 3), coincide on some μ-measurable set B with $E - B$ $\mathfrak{A}^*$-negligible (Theorem 4.3). If then A is an MT^*-space of Cauchy type, $A = \mathfrak{L}_C^\lambda = \Omega^A$

Received October 6, 1958.

on B. Thus an MT-space of Cauchy type on a locally compact space E that is countable at infinity coincides with a λ-space for μ on the restriction of E to some μ-measurable set B with $E - B$ $\mathfrak{A}^*$-negligible if and only if every element of A' is of base μ.

2. The MT-conjugates as vector spaces. Let E be a locally compact space, $\mathbf{C}^E$ the vector space of functions on E valued in $\mathbf{C}$ the field of complex numbers. A semi-norm on a vector subspace A of $\mathbf{C}^E$ will be called monotone if $\mathfrak{N}^A(x) \leqslant \mathfrak{N}^A(y)$ when $|x(t)| \leqslant |y(t)|$, $x, y \in A$; non-trivial if $\mathfrak{N}^A(x) \not\equiv 0$ over A **(6)**.

Definition. A vector subspace A of $\mathbf{C}^E$ will be called an MT^*-space if it contains $\mathfrak{K}_C$, if with x it contains $|x|$ and $\bar{x}$ and if it has a non-trivial, monotone semi-norm $\mathfrak{N}^A$.

If A' is the dual of A topologized by $\mathfrak{N}^A$ as a semi-norm and if $y \in A'$, then the restriction of y to $\mathfrak{K}_C$ determines a C-measure $\hat{y}$ and

$$y(x) = \int x \, d\hat{y}, \tag{2.1}$$

for every $x \in \mathfrak{K}_C$ **(6)**. We denote by A^* the subspace of elements y of A' for which every $x \in A$ is $\hat{y}$-integrable with (2.1) holding and call such a y an element of integral type. We call A^* the MT-conjugate of A. As in **(6)** the mapping $y \rightarrow \hat{y}$ of A^* into $\mathfrak{M}_C$, the space of measures on E, is an isomorphism. We denote by $\mathfrak{A}'$ and $\mathfrak{A}^*$ the images of A', A^* in $\mathfrak{M}_C$ and call $\mathfrak{A}^*$ the MT-measure conjugate of A. We define for each $y \in A^*$, $\hat{y} \in \mathfrak{A}^*$,

$$|\hat{y}|_{\mathfrak{A}^*} = \sup_{\substack{x \neq 0 \\ x \in A}} \left| \int x \, d\hat{y} \right| / \mathfrak{N}^A(x) = \sup_{\substack{x \neq 0 \\ x \in A}} |y(x)| / \mathfrak{N}^A(x) = |y|_{A^*} = |y|_{A'},$$

where $|y|_{A'}$ is the usual norm on A'. There are corresponding definitions for real MT^*-spaces.

A^* is a vector subspace of A'. Let $y_1, y_2 \in A^*$, $a, b \in C$. Then $z = ay_1 + by_2 \in A'$ and determines a C-measure $\hat{z}$. From (2.1) for $\mathfrak{K}_C$ it follows that $\hat{z} = a\hat{y}_1 + b\hat{y}_2$. By **(6**, Corollary 9.1**)** every $x \in A$ is $a\hat{y}_1 + b\hat{y}_2 = \hat{z}$-integrable and

$$\int x \, d\hat{z} = \int x \, d(a\hat{y}_1 + b\hat{y}_2) = ay_1(x) + by_2(x) = z(x).$$

The spaces A^* and $\mathfrak{A}^*$ are thus normed vector spaces, equivalent by definition.

Morse and Transue **(6**, p. 153**)** associate with each C-measure η on E a unique positive measure $|\eta|$ such that for $x \in K$, $x \geqslant 0$,

$$|\eta|(x) = \sup_{\substack{|u| \leqslant x \\ u \in \mathfrak{K}_C}} \left| \int u \, d\eta \right|. \tag{2.2}$$

The absolute measure $|\eta|$ defined by η then has a unique extension $|\eta|_e$ as a real C-measure on E **(6**, p. 151**)**.

Condition 2.1. *If $\eta \in \mathfrak{A}^*$, $|\eta|_e \in \mathfrak{A}^*$ and $|\eta|_{\mathfrak{A}^*} = ||\eta|_e|_{\mathfrak{A}^*}$.*

Condition (2.1) is the analogue for the MT-conjugate spaces of the condition for A that $|x| \in A$ if $x \in A$ (noting that the monotone property of $\mathfrak{N}^A$ implies that $\mathfrak{N}^A(x) = \mathfrak{N}^A(|x|)$). If, for a positive measure μ, the C-measure η is of base μ (that is, can be written in the form $g(t) \cdot \mu$ with $g(t)$ locally μ-integrable (**3**, p. 42; **7**, § 3).

$$(2.3) \qquad |g(t) \cdot \mu| = |g(t)| \cdot \mu.$$

When all the elements of $\mathfrak{A}^*$ are of base μ, A^* can be identified with the collection of functions $\{g(t)\}$. If then A^* is an MT^*-space Condition 2.1 is necessarily satisfied. We note also that it is trivially satisfied when $A^* = 0$, that it is satisfied by the measure dual of every MT-space (**6**, Lemma 11.2) and by the measure dual of every MT^*-space that is a λ-space with the MT- and λ-conjugates coinciding (**4**).

Suppose that α_i, $i = 1, 2, \ldots$, are positive measures with $\alpha_{i,e} \in \mathfrak{A}^*$ and that $\Sigma|\alpha_{i,e}|_{\mathfrak{A}^*} < \infty$. Then for every $x \in \mathfrak{K}$, $x > 0$,

$$\sum_1^\infty |\alpha_i(x)| = \sum_1^\infty \alpha_i(|x|) < \mathfrak{N}^A(x) \sum_1^\infty |\alpha_{i,e}|_{\mathfrak{A}^*} < \infty,$$

so that the α_i form a *summable family* of positive measures on E and determine a positive measure $\alpha_0 = \Sigma_1^\infty \alpha_i$ (**3**, § 3, no. 5).

THEOREM 2.1. *Let A be an MT^*-space for which Condition* 2.1 *holds. If every real x in A is α_0-integrable for every α_0 defined as in the preceding paragraph, then $\mathfrak{A}^*$ is complete.*

Proof. The theorem is trivial when $\mathfrak{A}^* = 0$. In the general case let $\{\eta_n\}$ denote a Cauchy sequence in $\mathfrak{A}^*$ and choose a subsequence $\{\eta_{n_i}\}$ with

$$|\eta_{n_1}|_{\mathfrak{A}^*} + \sum_1^\infty |\eta_{n_{i+1}} - \eta_{n_i}|_{\mathfrak{A}^*} = L < \infty.$$

Define

$$\alpha_1 = |\eta_{n_1}|, \alpha_i = |\eta_{n_{i+1}} - \eta_{n_i}|, i = 2, 3, \ldots, \alpha_0 = \sum_1^\infty \alpha_i.$$

Condition 2.1 implies that each $\alpha_{i,e}$ is in $\mathfrak{A}^*$ with

$$|\alpha_{1,e}|_{\mathfrak{A}^*} = |\eta_{n_1}|_{\mathfrak{A}^*},$$
$$|\alpha_{i,e}|_{\mathfrak{A}^*} = |\eta_{n_{i+1}} - \eta_{n_i}|_{\mathfrak{A}^*}, i = 1, 2, \ldots.$$

By hypothesis each real $x \in A$ is α_0-integrable so that (**3**, Proposition 5, 3°)

$$\int x \, d\alpha_0 = \sum_1^\infty \int x \, d\alpha_i.$$

If $x \in A$, $x = x_1 + ix_2$, with x_1 and x_2 real and in A, x is $\alpha_{0,e}$-integrable (**6**, Lemma 4.3) and

$$\int x\, d\alpha_{0,e} = \int x_1\, d\alpha_0 + i\int x_2\, d\alpha_0 = \sum_1^\infty \int x_1\, d\alpha_i + i\sum_1^\infty \int x_2\, d\alpha_i$$
$$= \sum_1^\infty \int x\, d\alpha_{i,e};$$
$$\left|\int x\, d\alpha_{0,e}\right| < \sum_1^\infty \int |x|\, d\alpha_{i,e} < L\,\mathfrak{N}^A(x).$$

It follows that $\int x\, d\alpha_{0,e}$ determines a continuous linear functional y of integral type with $\hat{y} = \alpha_{0,e}$ and therefore $\alpha_{0,e} \in \mathfrak{A}^*$.

For each $x \in A$,

$$[\eta_{n_{i+1}}(x) - \eta_{n_i}(x)]$$

is a Cauchy sequence in $\mathbf{C}$ since

$$\left|\sum_p^q [\eta_{n_{i+1}}(x) - \eta_{n_i}(x)]\right| < \sum_p^q \alpha_i(|x|) \to 0$$

as $p, q \to \infty$. Thus

$$\eta(x) = \eta_{n_1}(x) + \sum_1^\infty [\eta_{n_{1+i}}(x) - \eta_{n_i}(x)] = \lim_{i\to\infty} \eta_{n_i}(x) \tag{2.4}$$

is defined in $\mathbf{C}$ for every $x \in A$. Now η is linear on A and continuous since

$$|\eta(x)| < \alpha_0(|x|) < L\mathfrak{N}^A(x) \tag{2.5}$$

for all $x \in A$. Thus η determines an element of A'.

It follows from (2.2) and (2.5) that $|\eta|(x) < \alpha_0(x)$ for every $x > 0$, $x \in \mathfrak{R}$. This implies that $|\eta|^*(x) < \alpha^*{}_0(x)$ for every $x > 0$. Thus every α_0-negligible set is $|\eta|$-negligible and every α_0-measurable function is $|\eta|$-measurable (**2**, p. 180). Thus if $x \in A$, $|x|$ is $|\eta|$-measurable and x is η-measurable (**6**, p. 168). Since

$$\int |x| d\eta < \int |x| d\alpha_0 < L\,\mathfrak{N}^A(x) < \infty,$$

every x in A is η-integrable (**6**, Theorem 9.4). This with (2.5) shows that $\int x\, d\eta$ determines an element $y \in A^*$ with $\hat{y} = \eta$ so that $\eta \in \mathfrak{A}^*$.

Then

$$|\eta - \eta_{n_i}|_{\mathfrak{A}^*} = \sup_{0\neq x\in A} \left|\int x\, d(\eta - \eta_{n_i})\right|/\mathfrak{N}^A(x)$$
$$< \sup_{0\neq x\in A} \sum_{i+1}^\infty \int |x|\, d\alpha_j/\mathfrak{N}^A(x)$$
$$< \sum_{i+1}^\infty |\alpha_{j,e}|_{\mathfrak{A}^*}$$

which approaches zero as $i \to \infty$. The full sequence $\{\eta_n\}$ then converges to η in $\mathfrak{A}^*$ so that $\mathfrak{A}^*$ is complete.

COROLLARY. *If E is countable at infinity and A is an MT^*-space for which Condition* 2.1 *holds, A^* and $\mathfrak{A}^*$ are Banach spaces.*

Proof. By (**3**, Corollaire 2, p. 28) every $x \in A$ is α_0-integrable.

Length functions for a positive measure μ are defined in (**4**, **5**). We denote by $\mathfrak{L}^\lambda$, $\mathfrak{L}_C{}^\lambda$ the subspaces of $\mathbf{R}^E$ and $\mathbf{C}^E$ respectively consisting of μ-measurable functions $x(t)$ with $\lambda(x) = \lambda(|x|) < \infty$ (cf. **5**, p. 577). (If $x(t) \in \mathbf{C}^E$, it is μ-measurable for $\mu > 0$ if its Riesz components are μ-measurable (**6**. p. 168).)

We show that if $A = \bar{\mathfrak{L}}_C{}^1(E, \mu)$ (**4**, § 2) with E and μ defined as in (**2**, Exercise 4, pp. 116) A^* is not complete. We define $g_i(P) = 1/\ln n$, $P = (1/n, k/n^2)$, $n = 2, 3, \ldots i$; $g_i(P) = 0$ elsewhere; $g(P) = 1/\ln n$, $P = (1/n, k/n^2)$, $n = 2, 3, \ldots$; $g(P) = 0$ elsewhere. The g_i form a Cauchy sequence in A' and converge to g. Each $g_i \,.\, \mu$ is in $\mathfrak{A}^*$ but $g \,.\, \mu$ is not.

The λ-conjugate of every λ-space is complete since it is also a λ-space (**4**). Thus the MT-conjugate of an arbitrary λ-space containing $\mathfrak{K}_C$ is complete when it coincides with the λ-conjugate.

3. $\mathfrak{N}^A$-extensible MT*-spaces. For a normed or semi-normed space X we let X_u denote the subunit elements of X, that is, the elements with norm or semi-norm not exceeding unity (cf. **6**, p. 171).

Definition 3.1. A semi-norm $\mathfrak{N}^A$ on an MT^*-space A will be called *reflexive* if for every $x \in A$,

$$\mathfrak{N}^A(x) = \sup_{\eta \in \mathfrak{A}^*_u} |\textstyle\int x \, d\eta|. \tag{3.1}$$

THEOREM 3.1. *In order that $\mathfrak{N}^A$ be a reflexive semi-norm on the MT^*-space A it is necessary and sufficient that A^*_u be dense in A_u' for the $\sigma(A', A)$ topology.*

Proof. Since A_u' and A_u^* are *équilibré* parts of A', the polars of A_u' and A_u are respectively $A_u'^0 = (x \in A : |y(x)| \leqslant 1 \text{ for all } y \in A_u')$ and $A_u^{*0} = (x \in A : |y(x)| \leqslant 1 \text{ for all } y \in A^*_u)$ (**1**, p. 52). We first show that $A^{*0}_u = A_u'^0$. Since $A^*_u \subset A_u'$, $A_u'^0 \supset A^{*0}_u$ and it is sufficient to prove the opposite inequality. If $x \in A^{*0}_u$, the hypothesis that $\mathfrak{N}^A$ is reflexive implies that

$$\mathfrak{N}^A(x) = \sup_{y \in \mathfrak{A}^*_u} |\textstyle\int x \, d\hat{y}| \leqslant 1.$$

Thus $|y(x)| \leqslant \mathfrak{N}^A(x)|y|_{A'} \leqslant 1$ if $y \in A_u'$ so that $x \in A_u'^0$.

Thus $A^{*0}_u = A_u'^0$ and it follows that $A^{*00}_u = A_u'^{00} = A_u'$. Since A^*_u is convex and contains 0, the argument of (**1**, Proposition 3, p. 52) shows that $A_u' = A^{*00}_u$ is the closure of A^*_u for $\sigma(A', A)$.

We next prove that the condition is sufficient. Since the definition of $|y|_A{}^*$ implies that $\geqslant$ holds in (3.1) we need only show that, given $\epsilon > 0$, there exists $y \in A^*_u$ with $\mathfrak{N}^A(x) \leqslant |\int x \, dy| + \epsilon$.

By an extension of the Hahn-Banach Theorem there exsits $y_0 \in A_u'$ with $y_0(x) = \mathfrak{N}^A(x)$, $|y_0|_{A'} = 1$. The set $[y \in A'; |(y - y_0)(x)| < \epsilon]$ is a neighbourhood of y_0 for the $\sigma(A', A)$ topology and by hypothesis contains $y_1 \in A^*_u$. Then

$$0 \leqslant \mathfrak{N}^A(x) - |\textstyle\int x \, d\hat{y}_1| \leqslant |y_0(x) - y_1(x)| = |(y_0 - y_1)(x)| < \epsilon.$$

We note the analogy with the relation between E and E'' for Banach spaces (**1**, Proposition 5, p. 114).

Definition 3.2. A semi-norm on an MT^*-space will be called *extensible* if A satisfies Condition 2.1 and $\mathfrak{N}^A$ is reflexive. An MT^*-space will be called $\mathfrak{N}^A$-extensible if it has an extensible semi-norm.

For an extensible semi-norm

$$\mathfrak{N}^A(x) = \sup_{\eta \in \mathfrak{A}^*_u} \int^* |x| d|\eta| \tag{3.2}$$

holds with outer integrals replaced by integrals for every $x \in A$. Formula (3.2) then extends the definition of $\mathfrak{N}^A$ to all of $\mathbf{C}^E$ and all of $\bar{\mathbf{R}}^E$.

Given a collection of C-measures $\mathfrak{M}$ a function $x \in \mathbf{C}^E$ or $\bar{\mathbf{R}}^E$ will be called $\mathfrak{M}$-negligible if $|x(t)|$ is $|\eta|$-negligible for every $\eta \in \mathfrak{M}$. $\mathfrak{M}$-negligible sets, $\mathfrak{M}$-equivalence and almost everywhere ($\mathfrak{M}$) are then defined by analogy with the case where $\mathfrak{M}$ reduces to a single C-measure η. When A is an $\mathfrak{N}^A$-extensible MT^*-space, $\mathfrak{N}^A(x) = 0$ if x is $\mathfrak{A}^*$-negligible. If then $x(t)$ is defined and valued in $\mathbf{C}$ or $\bar{\mathbf{R}}$ almost everywhere ($\mathfrak{A}^*$), x is $\mathfrak{A}^*$-equivalent to some $\dot{x}$ in $\mathbf{C}^E$ or $\bar{\mathbf{R}}^E$ and we define $\mathfrak{N}^A(x) = \mathfrak{N}^A(\dot{x})$. When $\mathfrak{A}^* \equiv 0$ every function is $\mathfrak{A}^*$-negligible but $\mathfrak{N}^A(x) > 0$ holds for some $x \in A$.

THEOREM 3.2. *For $1 < p < \infty$, $A = \bar{\mathfrak{L}}_C{}^p(E, \mu)$ is an $\mathfrak{N}^A$-extensible MT^*-space.*

LEMMA 3.1. *If $A = \bar{\mathfrak{L}}_C{}^\lambda(E, \mu)$ is an MT^*-space for which the λ-conjugate contains the MT-conjugate, then Condition 2.1 is satisfied and every element of $\mathfrak{A}^*$ is of base μ.*

Proof of Lemma 3.1. Every g in the λ-conjugate is locally μ-integrable and therefore determines a measure $g \,.\, \mu$ (that is, a measure of base μ) (**4**, § 3). If $g \in A^*$, $\hat{g} = g \,.\, \mu$ and thus the elements of $\mathfrak{A}^*$ are of base μ.

The definition of the λ-conjugate then implies that $|g(t)| \in \mathfrak{L}^{\lambda*}$. By (**7**, § 3) $|g| \,.\, \mu = |g \,.\, \mu|$. Now $\int |x||g| d\mu \leqslant \lambda(x)\lambda^*(g) < \infty$ and the $g \,.\, \mu$-integrability of x implies that $\int |x| d|g \,.\, \mu| < \infty$ (**6**, Theorem 9.4). Thus by (**7**, Theorem 1.1), for every $x \in A$,

$$|g|(x) = \int x|g| d\mu = \int x \, d(|g| \,.\, \mu)_e$$

so that $(|g| \,.\, \mu)_e \in \mathfrak{A}^*$. It then follows from the definitions that

$$|(|g| \,.\, \mu)_e|_{\mathfrak{A}^*} = \lambda^*(|g|) = \lambda^*(g) = |g \,.\, \mu|_{\mathfrak{A}^*}.$$

Proof of Theorem 3.2. It remains to be shown that $\bar{\mathfrak{N}}^p$ is reflexive as a semi-norm on $A = \bar{\mathfrak{L}}_C{}^p$. Since $\bar{\mathfrak{N}}^p$ is reflexive as a length function,

$$\bar{\mathfrak{N}}^p(x) = \sup_{g \in (\mathfrak{L}_C^q)_u} |\int xg \, d\mu| \geqslant \sup_{g \in \mathfrak{A}^*_u} |\int xg \, d\mu|,$$

and it is sufficient to determine $g \in A^*_u$ with $|\int xg \, d\mu|$ arbitrarily near to $\bar{\mathfrak{N}}^p(x)$.

If $\mathfrak{N}^p(x) < \infty$ there exists $E_0 = \cup_1^\infty K_i$, where $\{K_i\}$ is an increasing sequence of compact sets for which, writing f_B for the product of the function $f(t)$ and the characteristic function of the set B,

$$\overline{\mathfrak{N}}^p(x) = \overline{\mathfrak{N}}^p(x_{E_0}) = \mathfrak{N}^p(x_{E_0})$$

(**4**, § 2). Now

$$x_{E_0} \in \mathfrak{L}_C{}^p$$

and $\mathfrak{L}_C{}^p$ is $\mathfrak{N}^A$-extensible as an MT-space. Thus

$$\overline{\mathfrak{N}}^p(x) = \mathfrak{N}^p(x_{E_0}) = \sup_{g \in (\mathfrak{L}^q_C)_u} |\int x_{E_0}\, g\, d\,\mu|.$$

Since E_0 is μ-measurable and

$$|g_{E_0}(t)| < |g(t)|, g_{E_0} \in \mathfrak{L}^q{}_u$$

if $g \in L_u{}^q$. Thus

$$\overline{\mathfrak{N}}^p(x) = \mathfrak{N}^p(x_{E_0}) = \sup_{g_{E_0} \in (\mathfrak{L}^q_C)_u} |\int x\, g_{E_0}\, d\,\mu|.$$

For $g \in (\mathfrak{L}_C{}^q)_u$ fixed,

$$|\int x\, g_{K_i}\, d\,\mu| \to |\int x\, g_{K_0}\, d\,\mu|$$

as $i \to \infty$ and

$$g_{K_i} \in (\mathfrak{L}^q_C)_u.$$

Thus for i sufficiently large and a suitable

$$g \in (\mathfrak{L}^q_C)_u, g_{K_i} \in (\mathfrak{L}^q_C)_u$$

with $|\int x\, g_{K_i} d\mu|$ arbitrarily near $\overline{\mathfrak{N}}^p(x)$. The C-measure $g_{K_i} \cdot \mu$ has compact support so that CK_i is $g_{K_i} \cdot \mu$-negligible (**2**, Proposition 5, p. 119). Thus if $f \in A$ and fg_{K_i} vanishes in E,

$$\int^* |f| d|g_{K_i} \cdot \mu| < \int^* |f|_{CK_i} d(|g_{K_i}| \cdot \mu) + \int^* |f|_{K_i} d(|g_{K_i}| \cdot \mu) = \int |fg_{K_i}| d\mu = 0$$

and the complex analogue of (**4**, Theorem 3.1) implies that

$$g_{K_i} \cdot \mu \in \mathfrak{A}_u{}^*.$$

THEOREM 3.3. *If λ is a reflexive length function for the positive measure μ, if E is countable at infinity or if E is arbitrary and $A = \mathfrak{L}_C{}^\lambda$ is an MT^*-space for which the MT- and λ-conjugates coincide, then A is $\mathfrak{N}^A$-extensible and every measure in $\mathfrak{A}^*$ is of base μ.*

Proof. Theorem 3.3 is a consequence of Lemma 3.1 and the fact that the reflexivity of $\mathfrak{N}^A = \lambda$ as a length function implies that it is reflexive as a semi-norm on A.

When A is an $\mathfrak{N}^A$-extensible MT^*-space we denote by $\mathfrak{F}^A$ the vector subspace of $\mathbf{C}^E$ of mappings x with $\mathfrak{N}^A(x) < \infty$. Then $\mathfrak{N}^A$ is a non-trivial,

monotone semi-norm on $\mathfrak{F}^A$ and $\mathfrak{F}^A$ is an MT^*-space for which Condition 2.1 holds. If for each $\eta \neq 0$ in $\mathfrak{A}^*$ there exists a relatively compact set $e(\eta)$ that is not η-measurable the MT-conjugate of $\mathfrak{F}^A$ reduces to the zero element of A'. Such non-measurable sets exist, for example, if $A = \mathfrak{L}_C^p\,(E, \mu)$ with $E = (0, 1)$ and μ Lebesgue measure on E, $1 < p < \infty$. In contrast, if E is arbitrary, if $A = \mathfrak{K}_C$ and $\mathfrak{N}^A$ is the uniform semi-norm, $\mathfrak{N}^A$ extends to $\mathbf{C}^E$ in the form **(6**, Theorem 15.3**)**,

$$\mathfrak{N}^A(x) = \sup_{t \in E} |x(t)|$$

and $\mathfrak{F}^A$ is the space of all bounded functions on E which is an $\mathfrak{N}^A$-extensible MT^*-space.

We note that if $B = \mathfrak{F}^A$, where A is an arbitrary $\mathfrak{N}^A$-extensible MT^*-space, $\mathfrak{N}^A(x) > 0$ is possible for a $\mathfrak{B}^*$-negligible function in B but $\mathfrak{N}^A(x) = 0$ for every $\mathfrak{A}^*$-negligible function in B.

The properties of the extended semi-norm $\mathfrak{N}^A$ and of $\mathfrak{F}^A$ for MT-spaces **(6**, § 12**)** extend to $\mathfrak{N}^A$-extensible MT^*-spaces with A'-negligibility replaced by $\mathfrak{A}^*$-negligibility. In particular $\mathfrak{F}^A$ is complete.

Generalizing **(6)** we define

$$\Omega^A = \bigcap_{\eta \in \mathfrak{A}^*} \mathfrak{L}_C^1(E, \eta)$$

for every MT^*-space A. We define $\Omega_0{}^A = \Omega^A \cap \mathfrak{F}^A$. Then $\Omega_0{}^A$ is an MT^*-space with $\mathfrak{N}^A$ (extended) as a semi-norm.

THEOREM 3.4. *If A is an $\mathfrak{N}^A$-extensible MT^*-space and if A^* is complete or, more generally, tonnelé* **(1**, § 1**)**, *then $\Omega_0{}^A = \Omega^A$.*

Proof. The argument of **(8**, Theorem 5.1**)** applies. We note in particular that $\Omega_0{}^A = \Omega^A$ for every $\mathfrak{N}^A$-extensible MT^*-space A if E is countable at infinity (Theorem 2.1, Corollary).

4. λ-spaces generated by $\mathfrak{N}^A$-extensible MT*-spaces.

THEOREM 4.1. *Let A be an $\mathfrak{N}^A$-extensible MT^*-space, μ a positive measure on E. Then $\mathfrak{N}^A$, extended by* (3.2), *defines a length function for μ if and only if every μ-negligible set is $\mathfrak{A}^*$-negligible.*

Proof. By (3.1) and the subsequent remarks $\mathfrak{N}^A(x)$ is defined for every $x(t)$ that is defined almost everywhere ($\mathfrak{A}^*$) and valued in $\overline{\mathbf{R}}^E$ and therefore for every $x(t)$, μ-measurable and defined, non-negative and valued in $\overline{\mathbf{R}}$ almost everywhere ($\mathfrak{A}^*$). That $\mathfrak{N}^A$ then satisfies Conditions (L2)–(L5) for length functions **(5)** is then easily verified. We verify (L5). If $x_n(t) \in \overline{\mathbf{R}}^E$ is non-negative and μ-measurable, $n = 1, 2, \ldots,$ and if $x_n(t)$ increases to $x(t)$ as $n \to \infty$, then for each $\eta \in \mathfrak{A}^*$,

$$\int^* x(t)\, d|\eta| = \sup_n \int^* x_n(t)\, d|\eta|,$$

by **(2**, Theorem 3, p. 110**)**. Thus

$$\mathfrak{N}^A(x) = \sup_{\eta \in \mathfrak{A}^*_u} \int^* x(t)\, d|\eta| = \sup_{\eta \in \mathfrak{A}^*_u} \sup_n \int^* x_n(t)\, d|\eta|$$
$$= \sup_n \mathfrak{N}^A(x_n).$$

If (L1) **(5)** holds every μ-negligible set is $\mathfrak{A}^*$-negligible. Conversely if every μ-negligible set is $\mathfrak{A}^*$-negligible, $\mathfrak{N}^A$ is defined and non-negative for every $x(t)$ that is non-negative a.e. (μ) (and therefore a.e. ($\mathfrak{A}^*$)) and if $x(t)$ is μ-negligible and $e = [t : x(t) \neq 0]$, e is μ-negligible **(2**, Theorem 1, p. 119**)** and therefore $\mathfrak{A}^*$-negligible. This implies that $x(t)$ is η-negligible for every $\eta \in \mathfrak{A}^*$ and (3.2) then shows that $\mathfrak{N}^A(x) = 0$ giving (L1).

We note that there exist $\mathfrak{N}^A$-extensible MT^*-spaces, in fact MT-spaces on a compact set E, for which $\mathfrak{N}^A$ cannot define a length function for any measure μ. Consider the MT-space $A = \mathfrak{C}_C(E)$ of complex valued functions continuous in $E = [0, 1]$ with semi-norm $\mathfrak{N}^A(x) = \sup_{t \in E}|x(t)|$ and suppose that $\mathfrak{N}^A$ defines a length function for some positive measure μ. Then, since $\mathfrak{A}^*$ contains all the point measures, the empty set is the only $\mathfrak{A}^*$-negligible set and therefore, by the preceding theorem, the only μ-negligible set. For each t, $0 < t < 1$, the set $\{t\}$ consisting of the point t is closed and therefore μ-measurable and $\mu(\{t\}) > 0$. For some $a > 0$ there is a collection of points t_i of E with $\mu(\{t_i\}) > a$, $i = 1, 2, \ldots$. Thus for the characteristic function of E, χ_E,

$$\mu(\chi_E) = \mu(E) > \lim_n \mu(\textstyle\bigcup_1^n t_i) > \lim_n na = \infty,$$

contradicting the assumption that μ is a measure since $\chi_E \in \mathfrak{C}_C$.

The following theorem is a partial converse of Theorem 3.3.

THEOREM 4.2. *Let A be an $\mathfrak{N}^A$-extensible MT^*-space, μ a positive measure on E and suppose that all of the elements of $\mathfrak{A}^*$ are of base μ. Suppose that every μ-negligible set is $\mathfrak{A}^*$-negligible and that every $\mathfrak{A}^*$-negligible set is locally μ-negligible. Then $A \subset \bar{A} \subset \mathfrak{L}_C{}^\lambda = \Omega_0{}^A \subset \mathfrak{F}^A$.*

Proof. By Theorem 4.1 $\mathfrak{N}^A$ determines a length function λ for μ. We denote by $\mathfrak{L}_C{}^\lambda$ the λ-space determined by λ. By hypothesis every $\eta \in \mathfrak{A}^*$ can be written $\eta = g \,.\, \mu$ where $g(t)$ is locally μ-integrable. We identify the functions $g(t)$ with A^*, the measures $g \,.\, \mu$ with $\mathfrak{A}^*$. If $E(g) = (t : g(t) \neq 0)$, $E(g)$ is μ-measurable and, for every $x \in \Omega^A$, $x_{E(g)}(t)$ is μ-measurable **(3**, Proposition 3, p. 43**)**. Given a compact set K in E with $\mu(K) > 0$ consider, for all $g \in A^*$, the collection of subsets $E(g)$ of K with $\mu[E(g)] > 0$. From this collection form a maximal collection of disjoint sets and let B denote their union. Since this collection will be at most countable B will be μ-measurable. If $g \in A^*$, $g_{K-B} \in A^*$ and $\mu[\mathrm{E}(g_{K-B})] = 0$ for otherwise $B \cup E(g_{K-B})$ properly contains B contradicting the definition of B. Thus, for every $g \in A^*$, $g(t) = 0$ almost everywhere in $K - B$, $g \,.\, \mu(K - B) = 0$ and $K - B$ is A^*-negligible and therefore, by hypothesis, $K - B$ is μ-negligible. If $x \in \Omega^A$, x_B is μ-measurable and therefore x_K is μ-measurable. It follows from **(2**, Proposition 4, p. 182**)**

that every $x \in \Omega^A$ is μ-measurable. If $x \in \Omega_0^A$, $\mathfrak{N}^A(x) < \infty$ and $x \in \mathfrak{L}_C^\lambda$. Thus $A \subset \Omega_0^A \subset \mathfrak{L}_C^\lambda$. Since $\mathfrak{L}_C^\lambda$ is complete it is closed in $\mathfrak{F}^A$ and contains $\bar{A}$, the closure of A.

To prove that $\mathfrak{L}_C^\lambda \subset \Omega_0^A$ we must show that every μ-measurable function $x(t)$ with $\mathfrak{N}^A(x) < \infty$ is in $\mathfrak{L}_C^1(g \,.\, \mu)$ for every $g \in A^*$. Every $x(t) \in \mathfrak{L}_C^\lambda$ is μ-measurable by definition so that the Riesz components of $x(t)$ are μ-measurable **(6**, p. 168**)**. The Riesz components are then measurable ($|g \,.\, \mu| = |g| \,.\, \mu$) for every $g \in A^*$ **(3**, Proposition 3, p. 43**)**. Thus $x(t)$ is measurable $(g \,.\, \mu)$ for every $g \in A^*$. Since for each $g \in A^*$, $|g \,.\, \mu|_e \in \mathfrak{A}^*$, it follows from (3.2) and **(6**, Theorem 9.4**)** that $x(t) \in \mathfrak{L}_C^1(g \,.\, \mu)$.

We note that if to each compact set K corresponds $g(t) \in A^*$ with $g(t) \neq 0$ a.e. (μ) in K, every $\mathfrak{A}^*$-negligible set is locally μ-negligible. This is true in particular if $\mathfrak{A}^*$ contains $\mathfrak{K}_C$ or the characteristic function of every compact set.

THEOREM 4.3. *Suppose that E is countable at infinity or that $E = E_0 \cup_1^\infty K_i$, with each K_i compact and E_0 locally μ-negligible, μ a positive measure. Let A be an $\mathfrak{N}^A$-extensible MT^*-space for which all of the elements of A^* are of base μ. Then, if E_0 is $\mathfrak{A}^*$-negligible, the normed spaces $\mathbf{L}_C^{\mathfrak{N}^A}$* and $\bar{\Omega}_0^A$ *associated with $\mathfrak{L}_C^{\mathfrak{N}^A}$ and Ω_0^A are equivalent and contain $\bar{A}$, the normed space associated with A.*

Proof. As in Theorem 4.2 each K_i is the union of a μ-measurable set B_i and an $\mathfrak{A}^*$-negligible set. If $B = \cup_1^\infty B_i$, x_B is μ-measurable for every $x \in \Omega^A$. Every $g \in A^*$ vanishes a.e. $(\mathfrak{A}^*)$ in $\cup_1^\infty K_i - B$. If not, for some g, i,

$$\mu[E(g_{K_i - B})] > 0,$$

contradicting the definition of B_i. It follows that $B' = E - B$ is $\mathfrak{A}^*$-negligible. Thus for each $x(t) \in \Omega^A$, $x_B(t)$ is μ-measurable and $\mathfrak{N}^A(x - x_B) = 0$. If then $x(t) \in \Omega_0^A$, $x_B(t) \in \mathfrak{L}_C^\lambda$, $\lambda = \mathfrak{N}^A$, with $\mathfrak{N}^A(x) = \mathfrak{N}^A(x_B)$ and $\bar{\Omega}_0^A \subset \mathbf{L}_C^\lambda$. The proof that $\mathbf{L}_C^\lambda \subset \bar{\Omega}_0^A$ is similar to the corresponding part of the proof of Theorem 4.2.

When E is countable at infinity $\Omega_0^A = \Omega^A$. The space E defined in **(2**, Exercise 4, p. 116**)** is of the form $D \cup_1^\infty K_i$ with D locally μ-negligible for the measure μ defined there. For the spaces $\bar{\mathfrak{L}}_C^p$, $1 < p < \infty$, D is $\mathfrak{A}^*$-negligible.

We note that if $\mathfrak{K}_C$ is dense in $\mathfrak{L}_C^\lambda$, $\bar{A} = \mathfrak{L}_C^\lambda$ in Theorem 4.2 and $\bar{\bar{A}} = \mathbf{L}_C^\lambda$ in Theorem 4.3.

5. MT*-spaces of Cauchy type. If A is an MT^*-space, let B be the vector subspace of A over R of real mappings in A, $\mathbf{B}$ the associated real normed vector space. As in **(8)**, with a natural definition of a partial order on $\mathbf{B}$, $\mathbf{B}$ becomes a "Riesz space."

Definition 5.1. A complete $\mathfrak{N}^A$-extensible MT^*-space will be called an MT^*-space of *Cauchy type* if each subset H of $\mathbf{B}$, bounded in norm and filtering for the relation $<$ defines a Cauchy filter.

For a maximal MT-space the definition reduces to that given in (**8**, § 1). The theory of MT-spaces of Cauchy type given in (**8**) extends to MT^*-spaces of Cauchy type with A'-negligibility replaced by $\mathfrak{A}^*$-negligibility and with Ω^A replaced by Ω_0^A.

THEOREM 5.1. *If A is an MT^*-space of Cauchy type then $A = \Omega_0^A$. If the hypotheses of Theorem 4.2 are then satisfied, $A = \mathfrak{L}_C^\lambda = \Omega_0^A$ and, if the hypotheses of Theorem 4.3 are satisfied $\bar{A} = \mathbf{L}_C^\lambda = \bar{\Omega}_0^A$.*

We note that if $A = \mathfrak{L}_C^\lambda$ is an MT^*-space of Cauchy type, the analogue of (**8**, Corollary 6.1) implies that λ satisfies (L9) (**4**, ((L9) as modified on p. 592)). Thus if $E = [0, 1]$, μ Lebesgue measure, the space $\mathfrak{L}_C^\infty (E, \mu)$ is not of Cauchy type.

REFERENCES

1. N. Bourbaki, *Éléments de Mathématique*, Fasc. XVIII, "Espaces Vectoriels Topologiques," chaps. III–V (Paris, 1955).
2. ——— *Éléments de Mathématique*, Fasc. XIII, "Integration," chaps. I–IV (Paris, 1952).
3. ——— *Éléments de Mathématique*, Fasc. XXI, "Integration," chap. V (Paris, 1956).
4. H. W. Ellis, *On the MT^*- and λ conjugates of $\mathfrak{L}^\lambda$* spaces, Can. J. Math., *10* (1958), 381–91.
5. H. W. Ellis and I. Halperin, *Function spaces determined by a levelling length function*, Can. J. Math., *5* (1953), 576–92.
6. Marston Morse and William Transue, *Semi-normed vector spaces with duals of integral type*, Jour. d'Analyse Math., *4* (1955), 149–86.
7. ——— *Products of a C-measure and a locally integrable mapping*, Can. J. Math., *9* (1957), 475–86.
8. ——— *Vector subspaces A of $\mathbf{C}^E$ with duals of integral type*, J. Math. pures et appl., Series 9, *37* (1958), 343–363.

Queen's University,
Summer Research Institute of the Canadian Mathematical Congress

ON A TYPE PROBLEM

JAMES A. JENKINS

Considerable interest has attached to the problem of determining the type of a Riemann surface obtained by performing an identification between the edges of a strip or a half-strip (**1**, **2**, **4**, **5**, **8**). A fairly thorough analysis was made in 1946 by Volkovyskii (**6**) who gave various sufficient conditions for parabolic and hyperbolic type. The object of the present paper is to show that his principal sufficient condition for hyperbolic type can be substantially improved.

We regard the half-strip S in the z-plane, $z = x + iy$

$$x > a, 0 < y < b$$

where a and b are finite real numbers, $b > 0$. (The case of a full strip is entirely equivalent.) We denote its edges $x > a$, $y = 0$ and $x > a$, $y = b$ respectively by L_1 and L_2. We consider the identification on L_1 and L_2 determined by the mapping (defined for $x \geqslant a$)

$$T(x) = f(x) + ib$$

where $f(x)$ is an increasing function of x with $f(a) > a$. We will suppose that this identification determines a Riemann surface $\mathscr{R}$ which is then doubly-connected with one boundary component C determined by the segments

$$x = a, 0 \leqslant y \leqslant b; a \leqslant x \leqslant f(a), y = b.$$

For this to hold it is necessary and sufficient that for each $x_0 > a$ there exist a disc $|w| < 1$ divided by a simple open arc λ into two domanins D_1 and D_2 and neighbourhoods E_1 and E_2 of x_0 and $T(x_0)$ relative to S such that there exist conformal mappings $w = \psi_i(z)$, $i = 1, 2$, of E_i onto D_i each admitting a homeomorphic extension to an open boundary arc γ_i of E_i and L_i which it carries onto λ and such that

$$\psi_1(x) = \psi_2(T(x)), \qquad x \in \gamma_1.$$

Non-trivial necessary and sufficient conditions on $f(x)$ for the identification T to determine a Riemann surface are not known. Some sufficient conditions were given by Volkovyskii (**7**). An easily verified sufficient condition is that $f(x)$ should possess a continuous derivative which does not take the value

Received August 20, 1958. Research supported in part by the National Science Foundation, Mathematics Section, through the University of Notre Dame and by the Office of Ordnance Research under contract No. DA-36-034-ORD-2453 through the Institute for Advanced Study.

zero. For our purposes here we assume that $f(x)$ and its inverse are absolutely continuous, that the identification T does determine a Riemann surface and that the extension of $\psi_i(z)$, $i = 1, 2$, to γ_i is continuously differentiable for each x_0 together with its inverse mapping. Without some restriction it is not *a priori* certain that such a surface is uniquely determined by the identification T. Volkovyskii **(7)** gave some sufficient conditions for this to hold but they are probably far from necessary.

A Riemann surface $\mathscr{R}$ determined by the above identification can be mapped conformally on the circular ring

$$1 < |\zeta| < R$$

where C corresponds to the boundary component $|\zeta| = 1$ and $R \leqslant \infty$. We will distinguish the cases according as $R < \infty$ or $R = \infty$ as being respectively hyperbolic or parabolic.

In order that we have the hyperbolic case it is sufficient that $\mathscr{R}$ have finite module for the family of curves Γ separating its boundary components (**3**, p. 13). That means that if $\rho\,(w)|dw|$ ranges over all conformally invariant metrics on $\mathscr{R}$ (w denotes a local uniformizing parameter) such that for any locally rectifiable simple closed curve γ separating the boundary components of $\mathscr{R}$

$$\int_\gamma \rho|dw| \geqslant 1$$

then ($w = u + iv$)

$$\min \iint_{\mathscr{R}} \rho^2\, du\, dv$$

is finite. In particular it is enough to manifest one admissible metric for which the above integral is finite. First, however, we wish to study the images in $\bar{S}$ of the family of curves Γ on $\mathscr{R}$.

To curves on $\mathscr{R}$ in Γ correspond sets on $\bar{S}$ which may display quite considerable complication. We will denote the family of such sets by Γ^*. Points of such a set γ^* on L_1 are accompanied by their images on L_2 under the mapping T. For any set ω in $\bar{S}$ we will call the (orthogonal) projection of $\omega - L_2$ on L_1 the projection $\pi(\omega)$ of ω. We denote the function $f(x)$ iterated n times by $f_n(x)$, also we take $f_0(x) \equiv x$. Points $(x', 0)$, $(x'', 0)$ on L_1 such that

$$x'' = f_n(x')$$

for some n will be called congruent points. The essential property of sets in Γ^* is given in the following lemma.

LEMMA. *For every $\gamma^* \in \Gamma^*$ there exists a value c, $c \geqslant a$, such that $\pi(\gamma^*)$ contains a point congruent to each point in the interval* $[c, f(c))$.

The set γ^* consists of arcs running from L_1 to L_2, intervals on L_1 and L_2 and arcs running from L_1 back to L_1 or from L_2 back to L_2. We note first that

since the corresponding curve $\gamma \in \Gamma$ is compact there can be only a finite number of arcs running from L_1 to L_2. Further we can replace these arcs by rectilinear segments with the same end-points not increasing their projections. If we replace an arc running from $L_1(L_2)$ back to $L_1(L_2)$ by the segment joining its end-points in the projection we at most replace a segment by a congruent segment. Finally a number of segments on $L_1(L_2)$ described consecutively (possibly overlapping) joining two points can be replaced by a single segment joining these points without increasing the projection. Thus it is enough to prove the result of the lemma when γ^* consists of a finite number of segments joining L_1 and L_2 and lying on L_1 and L_2.

There must be at least one segment joining L_1 and L_2. Let then P_1 be the end-point of such a segment farthest to the right on L_1 and P_2 be the end-point of such a segment farthest to the right on L_2. If $P_2 = T(P_1)$, γ^* consists of a single segment and the result of the lemma is evident. If P_2 is to the right of $T(P_1)$ it must be either the end-point of both a segment joining L_1 and L_2 and a segment on L_2 or the end-point of two segments joining L_1 and L_2. In the first instance replacing the segments by a segment forming with them a triangle (and deleting the corresponding segment on L_1) we obtain a new $\gamma' \in \Gamma^*$ with one less side (counting only one side for a pair of corresponding segments on L_1 and L_2) and a not larger projection. In the second instance replacing the segments by the segment on L_1 forming with them a triangle (and inserting the corresponding segment on L_2) we obtain a new $\gamma'' \in \Gamma^*$ with one less side and a not larger projection. Similarly if P_2 is to the left of $T(P_1)$, P_1 must be either the end-point of both a segment joining L_1 and L_2 and a segment on L_1 or the end-point of two segments joining L_1 and L_2. Proceeding as before the same conclusions apply. Since the result of the lemma is true for γ^* consisting of a single segment it follows in full generality by induction.

We now make our final assumption, that

$$\lim_{n\to\infty} f_n(a) = \infty. \tag{1}$$

We denote by I_n, $n = 0, 1, \ldots$, the region

$$f_n(a) \leq x < f_{n+1}(a),\ 0 \leq y < b.$$

Let $\phi_n(x)$ denote the function inverse to $f_n(x)$. Let $\mu(x)$ be a non-negative integrable function defined for $a \leq x < f(a)$ with

$$0 < \int_a^{f(a)} \mu(x)\, dx.$$

Then we consider in S the metric $\rho(z)|dz|$ where

$$\rho(z) = \mu(\phi_n(x))\phi_n'(x), \qquad z \in I_n.$$

Let $\gamma^* \in \Gamma^*$. Under assumption (1) it follows from our lemma that $\pi(\gamma^*)$ contains a point congruent to every point of the interval $[a, f(a))$. Thus

$$\int_{\gamma^*} \rho(z)|dz| > \int_a^{f(a)} \mu(x)\,dx.$$

On the other hand

$$\begin{aligned}\iint_S \rho^2(z)\,dx\,dy &= \sum_{n=0}^{\infty} \iint_{I_n} [\mu(\phi_n(x))\,\phi_n'(x)]^2\,dx\,dy \\ &= b \sum_{n=0}^{\infty} \int_{f_n(a)}^{f_{n+1}(a)} [\mu(\phi_n(x))\,\phi_n'(x)]^2\,dx \\ &= b \sum_{n=0}^{\infty} \int_a^{f(a)} \frac{(\mu(x))^2}{f_n'(x)}\,dx \\ &= b \int_a^{f(a)} (\mu(x))^2 \left(\sum_{n=0}^{\infty} \frac{1}{f_n'(x)}\right) dx\end{aligned}$$

provided that the operations involved are legitimate. This will be the case if

$$\sum_{n=0}^{\infty} \frac{1}{f_n'(x)}$$

converges at those points where $\mu(x)$ is positive and the last integral is convergent. In these circumstances the identification determined by f comes under the hyperbolic case.

We state our result as follows.

THEOREM. *If the function $f(x)$ defined for $a < x$ is absolutely continuous together with its inverse, if the identification it provides on the strip S determines a Riemann surface $\mathfrak{R}$, if the corresponding functions $\psi_i(z)$, $i = 1, 2$, admit extensions to the open boundary arcs of E_i on L_i which are continuously differentiable together with their inverse mappings, if $f(x)$ satisfies the condition*

$$\lim_{n\to\infty} f_n(a) = \infty \tag{1}$$

and if the series

$$\sum_{n=0}^{\infty} \frac{1}{f_n'(x)}$$

converges on a set of positive measure on $[a, f(a))$ then the identification comes under the hyperbolic case.

Indeed this sum is greater than or equal to one on the set of positive measure in question; thus we can take $\mu(x)$ in the preceding argument as the reciprocal of the sum on that set and elsewhere zero.

This result represents a substantial improvement of one of Volkovyskii's basic results which requires that the above series have a bounded sum on an interval as a sufficient condition for the hyperbolic case in addition to other requirements on the function $f(x)$ some of which are not germane to the present problem. It is immediately seen that the present considerations extend similarly to the case of identification of two strips also discussed by Volkovyskii **(6)**.

REFERENCES

1. C. Blanc, *Les surfaces de Riemann des fonctions méromorphes*, Comm. Math. Helvet., *9* (1936–7), 193–216, 335–68.
2. ———, *Les demi-surfaces de Riemann. Application au problème du type*, Comm. Math. Helvet., *10* (1938–9), 130–50.
3. James A. Jenkins, *Univalent functions and conformal mapping* (Springer-Verlag, Berlin-Göttingen-Heidelberg, 1958).
4. R. Nevanlinna, *Ueber die Polygondarstellung einer Riemannschen Fläche*, Ann. Acad. Sci. Fenn. Ser. AI, no. 122 (1952).
5. ———, "Polygonal representation of Riemann surfaces," *Lectures on functions of a complex variable* (Michigan, 1955), 65–70.
6. L. I. Volkovyskii, *On the problem of type of simply-connected Riemann surfaces*, Mat. Sbornik, *18* (N.S.) (1946), 185–211 (Russian, English summary).
7. ———, *Quasiconformal mapping and the problem of conformal pasting*, Ukrain. Mat. Z., *3* (1951), 39–52 (Russian).
8. E. M. Wirth, *Ueber die Bestimmung des Typus einer Riemannschen Fläche*, Comm. Math. Helvet., *31* (1956), 90–107.

University of Notre Dame
Institute for Advanced Study

ON SOME PROPERTIES OF FUNCTIONS ANALYTIC IN A HALF-PLANE

P. G. ROONEY

1. Introduction. The spaces $\mathfrak{H}_p(\omega)$, ω real, $1 < p < \infty$, consist of those functions $f(s)$, analytic for $\operatorname{Re} s > \omega$, and such that $\mu_p(f; x)$ is bounded for $x > \omega$, where

$$\mu_p(f; x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |f(x + iy)|^p \, dy. \tag{1.1}$$

Doetsch **(1)** has shown that if $e^{-\omega t}\phi(t) \in L_p\ (0, \infty)$, $1 < p < 2$, and f is the Laplace transform of ϕ, that is,

$$f(s) = \int_0^\infty e^{-st}\, \phi(t) dt, \qquad Re\, s > \omega,$$

then $f \in \mathfrak{H}_q(\omega)$, where

$$p^{-1} + q^{-1} = 1, \tag{1.2}$$

and that conversely if $f \in \mathfrak{H}_p(\omega)$, $1 < p < 2$, then there is a function ϕ, with $e^{-\omega t}\phi(t) \in L_q\ (0, \infty)$, such that f is the Laplace transform of ϕ.

The proofs of Doetsch's theorems are based on a generalization of Plancherel's theorem due to Titchmarsh **(5)**. Titchmarsh's theorem states that if $F \in L_p(-\infty, \infty)$, $1 < p < 2$, then F has a Fourier transform $G \in L_q\ (-\infty, \infty)$.

However, there are other extensions of Plancherel's theorem due to Hardy and Littlewood **(3)**. They have shown that if $F \in L_p\ (-\infty, \infty)$, $1 < p < 2$, then F has a Fourier transform G such that $|x|^{1-2/p}G(x) \in L_p\ (-\infty, \infty)$, and that conversely if $|x|^{1-2/q}F(x) \in L_q\ (-\infty, \infty)$, $q > 2$, then F has a Fourier transform $G \in L_q\ (-\infty, \infty)$—for this form of Hardy and Littlewood's theorems see **(7**, Theorems 79 and 80**)**. One might expect that a theory similar to Doetsch's theory could be constructed from these theorems, and this we shall do here.

To this end we define spaces $\mathscr{H}_p(\omega)$, $1 < p < \infty$, to consist of those functions $f(s)$ such that $(s - \omega)^{1-2/p}f(s) \in \mathfrak{H}_p(\omega)$ (where $(s - \omega)^{1-2/p}$ takes on its principal value). This is equivalent to saying that $\nu_p(f; x, \omega)$ should be bounded for $x > \omega$, where

$$\nu_p(f; x, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |x - \omega + iy|^{p-2} |f(x + iy)|^p \, dy. \tag{1.3}$$

In § 3 we shall obtain theorems corresponding to Doetsch's results for these new spaces. It will be noticed that $\mathfrak{H}_2(\omega) = \mathscr{H}_2(\omega)$, so that one would expect

Received July 28, 1958. This paper was written while the author was a fellow at the 1958 Summer Research Institute of the Canadian Mathematical Congress.

that our new theorems should reduce, for $p = 2$, to Doetsch's theorems. This is actually the case.

In an earlier paper **(4)** we generalized Doetsch's theory. In order to obtain theorems dealing with the Laplace transformation of functions, of the form $t^{\lambda}\phi(t)$, where $e^{-\omega t}\phi(t) \in L_p$ $(0, \infty)$ and $\lambda \geqslant 0$, we "generalized" the spaces $\mathfrak{H}_p(\omega)$ to spaces $\mathfrak{H}_{\lambda,p}(\omega)$. We can carry out a similar programme here, and to this end we define spaces $\mathscr{H}_{\lambda,p}(\omega)$ as follows. $\mathscr{H}_{0,p}(\omega) = \mathscr{H}_p(\omega)$; if $\lambda > 0$, $\mathscr{H}_{\lambda,p}(\omega)$ consists of those functions f in $\mathscr{H}_p(\omega')$ for every $\omega' > \omega$ such that $\nu_p^{\lambda}(f; \omega)$ is finite, where

$$\nu_p^{\lambda}(f; \omega) = \int_{\omega}^{\infty} (x - \omega)^{p\lambda-1} \ \nu_p(f; x, \omega)\, dx. \tag{1.4}$$

The theorems corresponding to the results of **(4)** are obtained in § 4.

In § 2 we prove certain preliminary lemmas concerning the properties of functions in $\mathscr{H}_p(\omega)$.

2. Preliminary lemmas.

LEMMA 1. *If* $f \in \mathscr{H}_p(\omega)$, $1 < p < \infty$, *then*

$$f(\omega + iy) = \lim_{x \to \omega+} f(x + iy)$$

exists for almost all y, *and* $|y|^{1-2/p} f(\omega + iy) \in L_p$ $(-\infty, \infty)$. *Further,* $(x - \omega + iy)^{1-2/p} f(x + iy)$ *converges in mean of order* p *to* $(iy)^{1-2/p} f(\omega + iy)$ *as* $x \to \omega+$. *Also,* ν_p $(f; x, \omega)$ *tends steadily from below, as* $x \to \omega+$, *to*

$$\int_{-\infty}^{\infty} |y|^{p-2} \, |f(\omega + iy)|^p \, dy.$$

Proof. The statement follows on applying **(1**, Lemma 7**)** to $F(z) = (z - \omega)^{1-2/p} f(z)$.

LEMMA 2. *Let* $f(s)$ *be analytic for* Re $s > \omega$, *and suppose*

$$\int_{-\infty}^{\infty} |x - \omega + iy|^{p-2} |f(x + iy)|^p \, dy$$

is bounded for $x_1 < x < x_2$, *where* $p > 1$, $x_1 > \omega$. *Then as* $y \to \pm\infty$, $f(x + iy) = o(|y|^{1-2/p})$, *uniformly in* x *for* $x_1 + \delta < x < x_2 - \delta$, *where* $0 < \delta < \frac{1}{2}(x_2 - x_1)$.

Proof. Let $\Phi(\zeta) = (-i\zeta)^{1-2/p} f(\omega - i\zeta)$, where $\zeta = \xi + i\eta$, and $(-i\zeta)^{1-2/p}$ has its principal value. Then if $\eta > 0$,

$$\int_{-\infty}^{\infty} |\Phi(\xi + i\eta)|^p \, d\xi = \int_{-\infty}^{\infty} |\eta - i\xi|^{p-2} |f(\omega + \eta - i\xi)|^p \, d\xi$$

$$= \int_{-\infty}^{\infty} |\eta + i\xi|^{p-2} \, |f(\omega + \eta + i\xi)|^p \, d\xi$$

which is bounded for $x_1 - \omega < \eta < x_2 - \omega$. Hence by **(7**, Lemma, p. 125**)**,

$\lim_{\xi\to\pm\infty}\Phi(\xi + i\eta) = 0$ uniformly in η for $x_1 - \omega + \delta < \eta < x_2 - \omega - \delta$. Thus, setting $x = \omega + \eta$, $y = -\xi$,

$$\lim_{y\to\pm\infty} (x - \omega + iy)^{1-2/p} f(x + iy) = 0$$

uniformly in x for $x_1 + \delta < x < x_2 - \delta$. But clearly

$$(x - \omega + iy)^{1-2/p} = O(|y|^{1-2/p}) \quad \text{as} \quad y \to \pm\infty,$$

uniformly in x for x in the same interval. Hence

$$\lim_{y\to\pm\infty} |y|^{1-2/p} f(x + iy) = 0$$

uniformly in x for x in this interval; that is,

$$f(x + iy) = o(|y|^{-(1-2/p)}) = o(|y|^{1-2/q})$$

uniformly in x for $x_1 + \delta < x < x_2 - \delta$.

LEMMA 3. *If* $f \in \mathscr{H}_q(\omega)$, $q > 2$, *and* $\omega < \xi < \operatorname{Re} s$, *then*

$$f(s) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{f(\xi + i\eta)}{s - (\xi + i\eta)}\,d\eta.$$

Proof. Suppose first $\omega < \xi < \operatorname{Re} s$. Let $s = x + iy$, and choose R and ρ so that $\rho > x$, and $R > |y|$. Then

$$f(s) = \frac{1}{2\pi i}\int\frac{f(\zeta)}{\zeta - s}\,d\zeta,$$

the integral being taken around the rectangle with vertices $\xi \pm iR$ and $\rho \pm iR$. The integral along the upper side of the rectangle is given by

$$\frac{1}{2\pi i}\int_{\xi}^{\rho}\frac{f(\alpha + iR)}{s - (\alpha + iR)}\,d\alpha.$$

But by Lemma 2, $f(\alpha + iR) = o(R^{1-2/p})$ as $R \to \infty$, uniformly in α for $\xi < \alpha < \rho$. Hence the integral along the upper side is $o(R^{-2/p})$ and consequently tends to zero as $R \to \infty$. Similarly, the integral along the lower side of the rectangle tends to zero as $R \to \infty$. Hence letting $R \to \infty$,

$$f(s) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{f(\xi + i\eta)}{s - (\xi + i\eta)}\,d\eta - \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{f(\rho + i\eta)}{s - (\rho + i\eta)}\,d\eta.$$

Now the second of these integrals tends to zero as $\rho \to \infty$. For from Hölder's inequality it is smaller in modulus than

$$(\nu_q(f; \rho, \omega))^{1/q}\left\{\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{|\rho - \omega + i\eta|^{p-2}}{|s - (\rho + i\eta)|^{p}}\,d\eta\right\}^{1/p}.$$

The first term of this expression is bounded by hypothesis; since $1 < p < 2$, the second term is smaller than

$$\left\{\frac{(\rho-\omega)^{p-2}}{2\pi}\int_{-\infty}^{\infty}\frac{d\eta}{((\rho-x)^2+(\eta-y)^2)^{1/2p}}\right\}^{1/p}$$

$$=\left\{\frac{(\rho-\omega)^{p-2}}{(\rho-x)^p}\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{d\eta}{(1+\eta^2)^{1/2p}}\right\}^{1/p}=O(\rho^{-2/p})$$

as $\rho\to\infty$. Hence letting $\rho\to\infty$

$$f(s)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{f(\xi+i\eta)}{s-(\xi+i\eta)}\,d\eta,\qquad \omega<\xi<Re\, s.$$

It remains to show that this equation remains true when $\xi=\omega$. For this we write the equation in the form

$$f(s)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\{(\xi-\omega+i\eta)^{1-2/q}f(\xi+i\eta)\}\left\{\frac{(\xi-\omega+i\eta)^{1-2/p}}{s-(\xi+i\eta)}\right\}d\eta.$$

The first term of the integrand of this last integral converges in mean of order q to $(i\eta)^{1-2/q}f(\omega+i\eta)$ as $\xi\to\omega+$. We shall show that the second term of the integrand converges in mean of order p to $(i\eta)^{1-2/p}/(s-(\omega+i\eta))$ as $\xi\to\omega+$. Clearly it tends to this limit pointwise. Further, since $1<p\leqslant 2$, we have if $\xi<\gamma<x$,

$$\left|\frac{(\xi-\omega+i\eta)^{1-2/p}}{s-(\xi+i\eta)}-\frac{(i\eta)^{1-2/p}}{s-(\omega+i\eta)}\right|^p$$

$$<2^p\left\{\frac{((\xi-\omega)^2+\eta^2)^{\frac{1}{2}(p-2)}}{((x-\xi)^2+(\eta-y)^2)^{\frac{1}{2}p}}+\frac{|\eta|^{p-2}}{((x-\omega)^2+(\eta-y)^2)^{1/2p}}\right\}$$

$$<2^{p+1}\cdot\frac{|\eta|^{p-2}}{((x-\gamma)^2+(\eta-y)^2)^{1/2p}}$$

which is in $L_1(-\infty,\infty)$ as a function of η. Hence by Lebesgue's theorem of dominated convergence,

$$\lim_{\xi\to\omega+}\int_{-\infty}^{\infty}\left|\frac{(\xi-\omega+i\eta)^{1-2/p}}{s-(\xi+i\eta)}-\frac{(i\eta)^{1-2/p}}{s-(\omega+i\eta)}\right|^p d\eta=0.$$

Thus, letting $\xi\to\omega+$ we obtain from **(6**, § 12.5, example (iv))**)**

$$f(s)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\{(i\eta)^{1-2/q}f(\omega+i\eta)\}\left\{\frac{(i\eta)^{1-2/p}}{s-(\omega+i\eta)}\right\}d\eta$$

$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{f(\omega+i\eta)}{s-(\omega+i\eta)}\,d\eta.$$

LEMMA 4. *If* $f\in\mathscr{H}_q(\omega)$, $q>2$, *and if* $\xi>\omega$ *and* $\mathrm{Re}\, s<\xi$, *then*

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{f(\xi+i\eta)}{s-(\xi+i\eta)}\,d\eta=0.$$

Proof. The statement follows much as in the previous lemma.

3. The spaces $\mathscr{H}_p(\omega)$. Theorems 1 and 2 correspond to Theorems 2 and 3 respectively of Doetsch **(1)**.

THEOREM 1. *If $e^{-\omega t}\phi(t) \in L_p\ (0, \infty)$, $1 < p < 2$, and*

$$f(s) = \int_0^\infty e^{-st}\,\phi(t)\,dt, \qquad Re\, s > \omega,$$

then $f \in \mathscr{H}_p(\omega)$ and if $x > \omega$,

$$\nu_p(f; x, \omega) < K \int_0^\infty e^{-pxt}\,|\phi(t)|^p\,dt,$$

where K depends on p alone.

Proof. If $x > \omega$,

$$f(x - iy) = \int_0^\infty e^{iyt}(e^{-xt}\phi(t))\,dt;$$

that is, for each fixed $x > \omega$, $f(x - iy)$ is the Fourier transform of a function in $L_p\ (0, \infty)$. Hence by **(7**, Theorem 80**)**, since $1 < p < 2$,

$$\begin{aligned}\nu_p(f; x, \omega) &= \frac{1}{2\pi}\int_{-\infty}^\infty |x - \omega + iy|^{p-2}\,|f(x + iy)|^p\,dy \\ &< \frac{1}{2\pi}\int_{-\infty}^\infty |y|^{p-2}|f(x - iy)|^p\,dy \\ &< \frac{K(p)}{2\pi}\int_0^\infty e^{-pxt}\,|\phi(t)|^p\,dt < \frac{K(p)}{2\pi}\int_0^\infty e^{-p\omega t}\,|\phi(t)|^p\,dt,\end{aligned}$$

so that $f \in \mathscr{H}_p(\omega)$ and the stated inequality holds with $K = K(p)/2\pi$.

THEOREM 2. *If $f \in \mathscr{H}_q(\omega)$, $q > 2$, then there is a function ϕ, with $e^{-\omega t}\phi(t) \in L_q(0, \infty)$, such that*

$$f(s) = \int_0^\infty e^{-st}\,\phi(t)\,dt, \qquad Re\, s > \omega.$$

Further, if $x > \omega$,

$$\int_0^\infty e^{-qxt}\,|\phi(t)|^q dt < K\,\nu_q(f; x, \omega),$$

where K depends on q alone.

Also for $x > \omega$ and for almost all t,

$$e^{xt}\,\underset{a\to\infty}{\mathfrak{L}_q}\,\frac{1}{2\pi}\int_{-a}^a e^{it\eta} f(x + i\eta)\,d\eta = \begin{cases}\phi(t), & t > 0\\ 0, & t < 0\end{cases}$$

(where $\mathfrak{L}_q$ denotes the limit in mean of order q).

Proof. By Lemma 1, $|y|^{1-2/q}f(\omega + iy) \in L_q(-\infty, \infty)$. Hence by **(7**, Theorem 79**)**, $f(\omega + iy)$ has a Fourier transform $F \in L_q\ (-\infty, \infty)$, given by the formula

$$F(t) = \underset{a\to\infty}{\mathfrak{L}_q} \frac{1}{(2\pi)^{\frac12}} \int_{-a}^{a} e^{it\eta} f(\omega + i\eta)\, d\eta.$$

Let $\phi(t) = (2\pi)^{-\frac12} e^{\omega t} F(t)$. Clearly $e^{-\omega t}\phi(t) \in L_q\ (-\infty, \infty)$.

Now for each s with $\operatorname{Re} s \neq \omega$, $(s - (\omega + i\eta))^{-1} \in L_p\ (-\infty, \infty)$ as a function of η. Also a straightforward calculation shows that if $\operatorname{Re} s > \omega$,

$$\frac{1}{(2\pi)^{\frac12}} (P) \int_{-\infty}^{\infty} \frac{e^{it\eta}}{s - (\omega + i\eta)}\, d\eta = \begin{cases} -\,(2\pi)^{\frac12}\, e^{t(s-\omega)}, & t > 0,\ Re\, s < \omega \\ (2\pi)^{\frac12}\, e^{t(s-\omega)}, & t < 0,\ Re\, s > \omega, \\ 0, & (Re\, s - \omega)\, t > 0, \end{cases}$$

so that the Fourier transform of $((s - (\omega + i\eta))^{-1}$ is given by this expression. Hence from Lemma 3 and (**7**, Theorem 81), if $\operatorname{Re} s > \omega$,

$$\begin{aligned} f(s) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{f(\omega + i\eta)}{s - (\omega + i\eta)}\, d\eta = \frac{1}{(2\pi)^{\frac12}} \int_{-\infty}^{0} e^{t(s-\omega)} F(-t)\, dt \\ &= \frac{1}{(2\pi)^{\frac12}} \int_{0}^{\infty} e^{-st} e^{\omega t} F(t)\, dt = \int_{0}^{\infty} e^{-st} \phi(t)\, dt, \end{aligned}$$

so that f is the Laplace transform of a function ϕ with $e^{-\omega t}\phi(t) \in L_q\ (0, \infty)$. Also from Lemma 4 and (**7**, Theorem 81), if $\operatorname{Re} s < \omega$,

$$\begin{aligned} 0 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{f(\omega + i\eta)}{s - (\omega + i\eta)}\, d\eta = -\,\frac{1}{(2\pi)^{\frac12}} \int_{0}^{\infty} e^{t(s-\omega)} F(-t)\, dt \\ &= -\,\frac{1}{(2\pi)^{\frac12}} \int_{0}^{\infty} e^{st} \phi(-t)\, dt, \end{aligned}$$

that is, the Laplace transform of $\phi(-t)$, with variable $-s$, vanishes. Hence by (**2**, chapter 2, § 9, Theorem 4) $\phi(-t) = 0$ a.e. for $t > 0$, or equivalently $\phi(t) = 0$ a.e. for $t < 0$.

Further, from (**7**, Theorem 79),

$$(3.1) \quad \int_{0}^{\infty} e^{-q\omega t} |\phi(t)|^q\, dt = \frac{1}{(2\pi)^{\frac12 q}} \int_{-\infty}^{\infty} |F(t)|^q\, dt$$
$$< \frac{K(q)}{2\pi} \int_{-\infty}^{\infty} |y|^{q-2}\, |f(\omega + iy)|^p\, dy.$$

Now since $q > 2$, if $\omega < \omega' < x$ and $g \in \mathscr{H}_q(\omega)$, then $\nu_q(g; x, \omega') < \nu_q(g; x, \omega)$, so that $g \in \mathscr{H}_q(\omega')$. Hence if $x > \omega$, $f \in \mathscr{H}_q(x)$ so that by what we have just proved there is a function ϕ_x with $e^{-xt}\phi_x(t) \in L_q\ (0, \infty)$, satisfying (3.1) with ω replaced by x, such that for $\operatorname{Re} s > x$,

$$f(s) = \int_{0}^{\infty} e^{-st} \phi_x(t)\, dt, \qquad Re\, s > x,$$

and so that for almost all t

$$e^{xt} \underset{a\to\infty}{\mathfrak{L}_q} \frac{1}{2\pi} \int_{-a}^{a} e^{it\eta} f(x + i\eta)\, d\eta = \begin{cases} \phi_x(t), & t > 0 \\ 0, & t < 0 \end{cases}$$

But by (**2**, chapter 2, § 9, Theorem 4), $\phi_x(t) = \phi(t)$ a.e. for $t > 0$. Hence for any $x > \omega$ and almost all t

$$e^{xt} \underset{a\to\infty}{\mathfrak{L}_q} \frac{1}{2\pi}\int_{-a}^{a} e^{it\eta} f(x+i\eta)\, d\eta = \begin{cases} \phi(t), & t > 0, \\ 0, & t < 0. \end{cases}$$

Finally from (3.1), with ω replaced by x, we obtain, since $q \geq 2$,

$$\int_0^\infty e^{-qxt}|\phi(t)|^q\, dt = \int_0^\infty e^{-qxt}|\phi_x(t)|^q\, dt \leq \frac{K(q)}{2\pi}\int_{-\infty}^{\infty} |y|^{q-2}|f(x+iy)|^q\, dy$$
$$\leq K\, \nu_q(f; x, \omega),$$

where $K = K(q)/2\pi$.

4. The spaces $\mathscr{H}_{\lambda,p}(\omega)$. Theorems 3 and 4 correspond to Theorems 1 and 2 of (**4**).

THEOREM 3. *If $e^{-\omega t}\phi(t) \in L_p\ (0, \infty)$, $1 < p \leq 2$, $\lambda \geq 0$, and*

$$f(s) = \int_0^\infty e^{-st}\, t^\lambda\, \phi(t)\, dt, \qquad Re\, s > \omega,$$

then $f \in \mathscr{H}_{\lambda,p}(\omega)$.

Proof. If $\lambda = 0$ the statement reduces to that of Theorem 1. Hence we may assume $\lambda > 0$. If $\omega' > \omega$, then since $t^\lambda e^{-(\omega'-\omega)t}$ is bounded for $t \geq 0$, $e^{-\omega' t}t^\lambda\phi(t) \in L_p\ (0, \infty)$, and hence by Theorem 1 $f \in \mathscr{H}_p(\omega')$, and if $x > \omega'$

$$\nu_p(f; x, \omega') \leq K\int_0^\infty e^{-pxt}\, t^{p\lambda}\, |\phi(t)|^p\, dt.$$

Let $x > \omega$, and choose ω' so that $\omega < \omega' < x$. Then since $1 < p \leq 2$,

$$\nu_p(f; x, \omega) \leq \nu_p(f; x, \omega') \leq K\int_0^\infty e^{-pxt}\, t^{p\lambda}\, |\phi(t)|^p\, dt.$$

Hence

$$\begin{aligned}
\nu_p^\lambda(f; \omega) &= \int_\omega^\infty (x-\omega)^{p\lambda-1}\, \nu_p(f; x, \omega)dx \\
&\leq K\int_\omega^\infty (x-\omega)^{p\lambda-1}\, dx \int_0^\infty e^{-pxt}\, t^{p\lambda}\, |\phi(t)|^p\, dt \\
&= K\int_0^\infty t^{p\lambda}\, |\phi(t)|^p\, dt \int_\omega^\infty (x-\omega)^{p\lambda-1}\, e^{-pxt}dt = \frac{K\Gamma(p\lambda)}{p^{p\lambda}}\int_0^\infty e^{-p\omega t}|\phi(t)|^p\, dt,
\end{aligned}$$

and $f \in \mathscr{H}_{\lambda,p}(\omega)$.

THEOREM 4. *If $f \in \mathscr{H}_{\lambda,q}(\omega)$, $q \geq 2$, $\lambda \geq 0$, then there is a function ϕ, with $e^{-\omega t}\phi(t) \in L_q\ (0, \infty)$, such that*

$$f(s) = \int_0^\infty e^{-st}\, t^\lambda\, \phi(t)\, dt.$$

Proof. Since $f \in \mathscr{H}_q(\omega')$ for every $\omega' > \omega$, by Theorem 2 if $\omega' > \omega$ there is a function $\phi_{\omega'}$, with

$$e^{-\omega' t}\, \phi_{\omega'}(t) \in L_q(0, \infty),$$

such that

$$f(s) = \int_0^\infty e^{-st}\, \phi_{\omega'}(t)dt, \qquad Re\, s > \omega'.$$

But by (**2**, chapter 2, § 9, Theorem 4), if ω' and ω'' are larger than ω, $\phi_{\omega'}(t) = \phi_{\omega''}(t)$ a.e. for $t > 0$. Hence if ϕ_0 is any one of these functions and Re $s > \omega$, then choosing ω' so that $\omega < \omega' <$ Re s we obtain

$$f(s) = \int_0^\infty e^{-st}\, \phi_{\omega'}(t)dt = \int_0^\infty e^{-st}\, \phi_0(t)dt.$$

Also from Theorem 2, since $q > 2$, if $x > \omega$ and ω' is chosen so that $\omega < \omega' < x$

$$\int_0^\infty e^{-qxt}\, |\phi_0(t)|^p dt = \int_0^\infty e^{-qxt}\, |\phi_{\omega'}(t)|^p dt < K\, \nu_q(f; x, \omega') < K\, \nu_q(f; x, \omega).$$

Hence, if we multiply this inequality by $(x - \omega)^{q\lambda-1}$ and integrate from ω to ∞, we obtain

$$\int_\omega^\infty (x - \omega)^{q\lambda-1}dx \int_0^\infty e^{-qxt}\, |\phi_0(t)|^q dt < K \int_\omega^\infty (x - \omega)^{q\lambda-1}\, \nu_q(f; x, \omega)dx$$
$$= K\nu_q^\lambda(f; \omega).$$

But the integral on the left-hand side of this inequality is equal to

$$\int_\omega^\infty (x - \omega)^{q\lambda-1}dx \int_0^\infty e^{-qxt}\, |\phi_0(t)|^q dt = \int_0^\infty |\phi_0(t)|^q dt \int_\omega^\infty (x - \omega)^{q\lambda-1} e^{-qxt} dx$$
$$= \frac{\Gamma(q\lambda)}{q^{q\lambda}} \int_0^\infty e^{-q\omega t}\, t^{-q\lambda}\, |\phi_0(t)|^q\, dt,$$

so that

$$\int_0^\infty e^{-q\omega t}\, t^{-q\lambda}\, |\phi_0(t)|^q dt < \frac{q^{q\lambda} K \nu_q(f, \omega)}{\Gamma(q\lambda)} < \infty.$$

Hence if we let $\phi(t) = t^{-\lambda}\phi_0(t)$, then $e^{-\omega t}\phi(t) \in L_q\ (0, \infty)$, and if Re $s > \omega$

$$f(s) = \int_0^\infty e^{-st}\, t^\lambda\, \phi(t)dt.$$

References

1. G. Doetsch, *Bedingungen für die Darstellbarkeit einer Funktion als Laplace-Integral und eine Umkehrformel für die Laplace-Transformation*, Math. Zeit., *42* (1937), 263–86.
2. ———, *Handbuch der Laplace-Transformation. I* (Basel, 1950).
3. G. H. Hardy and J. E. Littlewood, *Some new properties of Fourier constants*, Math. Annal., *97* (1926), 159–209.
4. P. G. Rooney, *On some theorems of Doetsch*, Can. J. Math., *10* (1958), 421–30.
5. E. C. Titchmarsh, *A contribution to the theory of Fourier transforms*, Proc. Lond. Math. Soc. (2), *23* (1923), 279–89.
6. ———, *The theory of functions* (2nd ed.; Oxford, 1939).
7. ———, *Introduction to the theory of Fourier integrals* (2nd ed.; Oxford, 1948).

University of Toronto

A NETWORK-FLOW FEASIBILITY THEOREM AND COMBINATORIAL APPLICATIONS

D. R. FULKERSON

1. Introduction. There are a number of interesting theorems, relative to capacitated networks, that give necessary and sufficient conditions for the existence of flows satisfying constraints of various kinds. Typical of these are the supply-demand theorem due to Gale **(4)**, which states a condition for the existence of a flow satisfying demands at certain nodes from supplies at other nodes, and the Hoffman circulation theorem (received by the present author in private communication), which states a condition for the existence of a circulatory flow in a network in which each arc has associated with it not only an upper bound for the arc flow, but a lower bound as well. If the constraints on flows are integral (for example, if the bounds on arc flows for the circulation theorem are integers), it is also true that integral flows meeting the requirements exist provided any flow does so. This fact has been used by Gale **(4)**, and by Ford and Fulkerson **(3)**, in the solution of several combinatorial problems. For example, Gale has shown how the supply-demand theorem, together with the existence of integral flows, can be used to derive simple conditions for the existence of a matrix of zeros and ones having prescribed row and column sums, a problem that was also solved independently by Ryser **(9)** by means of purely combinatorial methods.

The present paper adds some results along the lines we have described. We first establish a feasibility theorem, which may be described informally as follows. Suppose there is given a capacitated network with certain of the nodes designated as sources, others as sinks, and assume that each source is required to send, and each sink to receive, an amount that lies between prescribed bounds. Under what conditions is this possible? The theorem asserts that if (a) there is a flow that sends out of each source an amount at least as great as the lower bound for the source, and into each sink no more than the upper bound for the sink, and if (b) there is a flow that sends out of each source no more than the upper bound for the source, and into each sink at least as much as the lower bound for the sink, then there is a flow that meets all the requirements simultaneously. We do not give a direct proof of this theorem, but rather use the max-flow min-cut theorem **(1; 2)** to find a pair of conditions that are necessary and sufficient for the existence of the required flow, and then observe that one of the conditions is equivalent to (a) above, the other to (b).

Received July 28, 1958.

Our first combinatorial application (§ 5) of the feasibility theorem is to generalize the Gale-Ryser theorem on incidence matrices having prescribed row and column sums, to the extent of allowing these sums to vary within designated bounds.

Our second application (§ 6) concerns the subgraph problem for directed graphs: to find necessary and sufficient conditions that a finite directed graph G have a subgraph H possessing specified local degrees. A solution to this problem has been given by Ore **(6)**. Here, in keeping with the feasibility theorem, we extend the problem by permitting the number of arcs of H that enter or leave each node of G to vary within bounds, and then show that the conditions obtained for this latter problem reduce to Ore's conditions for the subgraph problem.

The similar problem for undirected graphs, which has been solved by Tutte **(12)**, and also by Ore **(8)**, is, so far as we known, amenable to network-flow methods only in the special case that G is an even graph, and this because the problem then is, in essence, a directed one.

Our final application (§ 7) deals with a problem involving set representatives: to find necessary and sufficient conditions for the existence of a system of distinct representatives having the further property that the intersection of the system with each member of a given partition of the fundamental set has a cardinality lying between assigned bounds. This problem was first posed and solved by Hoffman and Kuhn **(5)**; it is shown here that the conditions established in **(5)** are deducible from the feasibility theorem.

2. Definitions, notation, and prior results. Let G be a finite directed network or linear graph consisting of a set N of nodes, $x, y, \ldots$, and directed arcs joining pairs of nodes, the arc from x to y being denoted (x, y), and suppose that each arc (x, y) has associated with it a *capacity* $c(x, y)$, where $c(x, y)$ is either a non-negative real number or plus infinity. Let the set N of nodes be partitioned into three subsets: S (the set of *sources*), T (the set of *sinks*), and R (the set of *intermediate nodes*). We call a real-valued function f defined on the arcs of G a *flow from S to T* provided that

$$(1) \qquad \sum_{y \in A(x)} f(x, y) = \sum_{y \in B(x)} f(y, x), \qquad x \in R,$$

$$(2) \qquad 0 \leq f(x, y) \leq c(x, y), \qquad \text{all } (x, y),$$

where $A(x)$ ("after" x) is the set of nodes y such that (x, y) is an arc, and $B(x)$ ("before" x) consists of those nodes y such that (y, x) is an arc. Thus (1) states that the flow out of an intermediate node is equal to the flow in, and (2) that the flow in each arc does not exceed its capacity.

We will be interested in flows from S to T that satisfy bounds on the net flow leaving each $x \in S$, and entering each $x \in T$. Thus, for $x \in S$, let $\alpha(x)$ and $\beta(x)$ be real-valued functions with

$$0 \leq \alpha(x) \leq \beta(x);$$

similarly, associate with each $x \in T$ two real numbers $a(x)$ and $b(x)$, where

$$0 \leq a(x) \leq b(x).$$

The additional constraints

$$\text{(3a)} \qquad \alpha(x) \leq \sum_{y \in A(x)} f(x, y) - \sum_{y \in B(x)} f(y, x) \leq \beta(x), \qquad x \in S,$$

$$\text{(3b)} \qquad a(x) \leq \sum_{y \in B(x)} f(y, x) - \sum_{y \in A(x)} f(x, y) \leq b(x), \qquad x \in T,$$

will be termed *feasible* provided there is a flow f from S to T satisfying them. In this case, f will also be called a *feasible flow*.

To simplify the notation, we adopt the following conventions. If X and Y are subsets of N, denote by (X, Y) the set of arcs leading from X to Y; and for any fuction f defined on the arcs, let

$$\sum_{(x,y) \in (X,Y)} f(x, y) = f(X, Y).$$

Similarly, if α is defined on a subset X of N, let

$$\sum_{x \in X} \alpha(x) = \alpha(X).$$

We shall also use $A(X)$ to denote the set of all nodes y such that (x, y), for some $x \in X$, is an arc of G, and similarly for $B(X)$.

The *value* $v(f)$ of a flow f from S to T is the net flow leaving the sources, which, in the notation just introduced, is given by

$$\text{(4)} \qquad v(f) = f(S, A(S)) - f(B(S), S).$$

In view of (1), $v(f)$ may also be expressed as the net flow entering the sinks:

$$\text{(5)} \qquad v(f) = f(B(T), T) - f(T, A(T)).$$

Let $X, \bar{X}$ be a partition of N with $S \subset X$, $T \subset \bar{X}$. The set of arcs $(X, \bar{X})$ is a *cut* in G (separating S and T), and $c(X, \bar{X})$ is the *cut capacity*.

A fundamental theorem concerning flows from S to T in a network G asserts that the maximal flow value is equal to the minimal cut capacity **(1, 2)**. A second theorem, important for combinatorial applications, is that if the capacity function c assumes only integral vlaues, then there exists a maximal flow f that is likewise integral **(2, 3)**.

Gale **(4)** has used the max-flow min-cut theorem to prove that if $\alpha(x) = 0$, $b(x) = \infty$ in (3), then a feasible flow (that is, a flow satisfying the "demands" $a(x)$ at the sinks from the "supplies" $\beta(x)$ at the sources) exists if and only if, for every partition $X, \bar{X}$ of N, we have

$$\text{(6)} \qquad a(T \cdot \bar{X}) \leq c(X, \bar{X}) + \beta(S \cdot \bar{X}),$$

where $X \cdot Y$ denotes the intersection of the sets X and Y.

3. Feasibility theorems. In this section, we develop a generalization of the supply-demand feasibility theorem by finding conditions under which the full set of constraints (3) is feasible.

We begin by adjoining to the given network G four new nodes, s, t, u, v, and several sets of arcs, as follows:

$$(s, S),\ (u, S),\ (T, t),\ (T, v),\ (u, t),\ (s, v),\ (t, s).$$

Next, we extend the capacity function c defined on arcs of G to the new network G^* by

$$\begin{aligned} c(s, x) &= \beta(x) - \alpha(x), && x \in S,\\ c(u, x) &= \alpha(x), && x \in S,\\ c(x, t) &= b(x) - a(x), && x \in T,\\ c(x, v) &= a(x), && x \in T,\\ c(u, t) &= a(T), &&\\ c(s, v) &= \alpha(S), &&\\ c(t, s) &= \infty. && \end{aligned}$$

We assert that a feasible flow exists in G if, and only if, the value of a maximal flow from u to v in G^* is $\alpha(S) + a(T)$. Suppose first that f is feasible in G; extend f to f^*, defined on the arcs of G^*, as follows:

$$\begin{aligned} f^*(s, x) &= f(x, A(x)) - f(B(x), x) - \alpha(x), && x \in S,\\ f^*(u, x) &= \alpha(x), && x \in S,\\ f^*(x, t) &= f(B(x), x) - f(x, A(x)) - a(x), && x \in T,\\ f^*(x, v) &= a(x), && x \in T,\\ f^*(u, t) &= a(T), &&\\ f^*(s, v) &= \alpha(S), &&\\ f^*(t, s) &= f(S, A(S)) - f(B(S), S), &&\\ f^*(x, y) &= f(x, y), && \text{for arcs } (x, y) \text{ of } G. \end{aligned}$$

It is a routine matter to check that f^* is a flow from u to v in G^*. Clearly, f^* has value

$$v(f^*) = \alpha(S) + a(T).$$

Conversely, let f^* be a flow from u to v in G^*, of value $\alpha(S) + a(T)$. Then

$$\begin{aligned} f^*(u, x) &= \alpha(x), && x \in S,\\ f^*(x, v) &= a(x), && x \in T. \end{aligned}$$

Let f be f^* restricted to G. Then f is a flow from S to T in G, and it remains only to show that f is feasible. Consider any $x \in S$. From (1) applied to x, we have

$$f^*(u, x) + f^*(s, x) = f(x, A(x)) - f(B(x), x),$$

or

$$\alpha(x) + f^*(s, x) = f(x, A(x)) - f(B(x), x);$$

and, since

$$0 \leqslant f^*(s, x) \leqslant \beta(x) - \alpha(x),$$

we get

$$\alpha(x) \leqslant f(x, A(x)) - f(B(x), x) \leqslant \beta(x),$$

which is (3a). Inequalities (3b) are similarly proved. This completes the proof of the assertion.

We may, therefore, in searching for feasibility criteria, rephrase the question as follows. Under what conditions does there exist a flow f^* from u to v in G^* having value $v(f^*) = \alpha(S) + a(T)$—that is, saturating all source and sink arcs? The max-flow min-cut theorem can now be used to provide an answer to this question by insisting that the capacitities of all cuts separating u and v be at least as great as $\alpha(S) + a(T)$.

Thus, let $(X^*, \bar{X}^*)$ be a cut in G^*, and consider cases.

Case 1. $s \in X^*$, $t \in \bar{X}^*$. Partition $X^*, \bar{X}^*$ as follows: $X^* = u + s + X$, $\bar{X}^* = v + t + \bar{X}$. Then

$$\begin{aligned} c(X^*, \bar{X}^*) &= c(u, t) + c(u, \bar{X}), + c(s, v) + c(s, \bar{X}) \\ &\quad + c(X, v) + c(X, t) + c(X, \bar{X}) \\ &= a(T) + \alpha(S \cdot \bar{X}) + \alpha(S) + \beta(S \cdot \bar{X}) - \alpha(S \cdot \bar{X}) \\ &\quad + a(T \cdot X) + b(T \cdot X) - a(T \cdot X) + c(X, \bar{X}). \end{aligned}$$

Hence, in this case, we always have $c(X^*, \bar{X}^*) \geqslant \alpha(S) + a(T)$.

Case 2. $s \in \bar{X}^*$, $t \in X^*$. Then $c(X, \bar{X}^*)$ is infinite. Hence again no condition is obtained.

Case 3. $s \in X^*$, $t \in X^*$. Letting $X^* = s + t + u + X$, $\bar{X}^* = v + \bar{X}$, we have

$$\begin{aligned} c(X^*, \bar{X}^*) &= c(s, v) + c(s, \bar{X}) + c(u, \bar{X}) + c(X, v) + c(X, \bar{X}) \\ &= \alpha(S) + \beta(S \cdot \bar{X}) - \alpha(S \cdot \bar{X}) + \alpha(S \cdot \bar{X}) \\ &\quad + a(T \cdot X) + c(X, \bar{X}). \end{aligned}$$

Thus $c(X^*, \bar{X}^*) \geqslant \alpha(S) + a(T)$ if, and only if,

$$\beta(S \cdot \bar{X}) + c(X, \bar{X}) \geqslant a(T \cdot \bar{X}). \tag{7}$$

Case 4. $s \in \bar{X}^*$, $t \in \bar{X}^*$. Let $X^* = u + X$, $\bar{X}^* = s + t + v + \bar{X}$. Then

$$\begin{aligned} c(X^*, \bar{X}^*) &= c(u, t) + c(u, \bar{X}) + c(X, t) + c(X, v) + c(X, \bar{X}) \\ &= a(T) + \alpha(S \cdot \bar{X}) + b(T \cdot X) - a(T \cdot X) \\ &\quad + a(T \cdot X) + c(X, \bar{X}), \end{aligned}$$

and we obtain the condition

$$b(T \cdot X) + c(X, \bar{X}) \geqslant \alpha(S \cdot X). \tag{8}$$

We may therefore state the following result.

THEOREM 1. *The constraints* (3) *are feasible if and only if* (7) *and* (8) *hold for all partitions* $X, \bar{X}$ *of* N.

Notice that (7) is precisely condition (6) for the supply-demand case; that is, if $\alpha(x) = 0$ for $x \in S$, and $b(x) = \infty$ for $x \in T$, then Theorem 1 reduces to the supply-demand theorem of **(4)**. Condition (8) may be interpreted as follows. If we interchange sources and sinks in G, reverse all arc directions, and think of α as the demand function at the set S of sinks, b as the supply function at the set T of sources, then (8) is a necessary and sufficient condition for feasibility of the supplies and demands in the reversed network. Thus Theorem 1 may be restated as follows.

THEOREM 2. *The constraints* (3a) *and* (3b) *are jointly feasible if, and only if, the constraints*

$$\begin{cases} \alpha(x) \leqslant f(x, A(x)) - f(B(x), x), & x \in S, \\ f(B(x), x) - f(x, A(x)) \leqslant b(x), & x \in T, \end{cases} \tag{9}$$

and

$$\begin{cases} f(x, A(x)) - f(B(x), x) \leqslant \beta(x), & x \in S, \\ a(x) \leqslant f(B(x), x) - f(x, A(x)), & x \in T, \end{cases} \tag{10}$$

are separately feasible.

Theorem 2 is the formulation described verbally in the Introduction. One suspects that there should be a simple method of constructing a flow satisfying all the constraints from the two separate flows, but we have not found such a method.

We note one other fact for the combinatorial applications. Namely, if the functions α, β, a, b, and c are integral-valued, and if the constraints (3) are feasible, then there is an integral feasible flow f. This follows directly from the proof of Theorem 1 and the existence of integral maximal flows in networks having integral capacities.

4. Application to matrices. When the network G is suitably specialized, Theorem 2 (or Theorem 1) provides criteria for the existence of a non-negative matrix whose row and column sums lie between designated limits, or, more generally, for the existence of a matrix with this property and the further property that the elements of the matrix are bounded above by specified numbers. We state the criteria provided by Theorem 2 explicitly as follows:

THEOREM 3. *Let* $0 \leqslant \alpha_i \leqslant \beta_i$, $i = 1, \ldots, m$, $0 \leqslant a_j \leqslant b_j$, $j = 1, \ldots, n$, *and* $c_{ij} \geqslant 0$ *be given constants. If there are matrices* f^1_{ij}, f^2_{ij} *satisfying*

$$\alpha_i \leqslant \sum_j f^1_{ij}, \qquad \sum_i f^1_{ij} \leqslant b_j, \qquad 0 \leqslant f^1_{ij} \leqslant c_{ij}, \tag{11}$$

$$\sum_j f^2_{ij} \leqslant \beta_i, \qquad a_j \leqslant \sum_i f^2_{ij}, \qquad 0 \leqslant f^2_{ij} \leqslant c_{ij}, \tag{12}$$

then there is a matrix f_{ij} *satisfying*

$$(13)\qquad \alpha_i \leq \sum_j f_{ij} \leq \beta_i, \qquad a_j \leq \sum_i f_{ij} \leq b_j, \qquad 0 \leq f_{ij} \leq c_{ij}.$$

To prove Theorem 3, take G to be the network consisting of nodes x_i $(i = 1, \ldots, m)$, y_j $(j = 1, \ldots, n)$, and arcs (x_i, y_j) of capacity c_{ij}. Let $S = \{x_1, \ldots, x_m\}$, $T = \{y_1, \ldots, y_n\}$, so that R is vacuous. Associate with each source x_i the bounds α_i, β_i, and with each sink y_j the bounds a_j, b_j. Then a flow from S to T is a matrix f_{ij} satisfying $0 \leq f_{ij} \leq c_{ij}$; a feasible flow satisfies, in addition, the first two inequalities of (13). Thus Theorem 3 is a direct consequence of Theorem 2.

5. Incidence matrices. Gale **(4)** and Ryser **(9)** have found simple conditions for the existence of a matrix of zeros and ones having prescribed row and column sums—or, what is the same thing, for the existence of an incidence matrix whose row sums are bounded below by given integers and whose column sums are bounded above by given integers.

The following is one interpretation of their problem. Suppose there is given a finite set $E = \{e_1, \ldots, e_m\}$. Under what conditions on the sets of integers $\{\alpha_1, \ldots, \alpha_m\}$ and $\{b_1, \ldots, b_n\}$ is it possible to construct n subsets $E_1, \ldots, E_n$ of E such that (a) the number of sets E_j that contain the element e_i is at least α_i, and (b) the set E_j contains at most b_j elements?

The conditions are surprisingly simple. Arrange the α's in decreasing order,

$$\alpha_{i_1} \geq \alpha_{i_2} \geq \ldots \geq \alpha_{i_m},$$

and define σ_k to be the number of integers in the set of b's that are greater than or equal to k. Then the required incidence matrix exists if, and only if, we have

$$(14)\qquad \sum_{k=1}^{l} \alpha_{i_k} \leq \sum_{k=1}^{l} \sigma_k, \qquad l = 1, 2, \ldots,$$

where we take $\alpha_{i_k} = 0$ for $k > m$.

As a corollary of the Gale-Ryser condition (14), Theorem 3 with all $c_{ij} = 1$, and the remark at the end of § 3, we have the following result:

THEOREM 4. *There exists a matrix of zeros and ones for which the* ith *row sum lies between given non-negative integers* α_i *and* β_i, *and the* jth *column sum lies between given non-negative integers* a_j *and* b_j, *where* $\alpha_i \leq \beta_i$, $a_j \leq b_j$, *if, and only if,*

$$(15)\qquad \sum_{k=1}^{l} \alpha_{i_k} \leq \sum_{k=1}^{l} \sigma_k, \qquad l = 1, 2, \ldots,$$

$$(16)\qquad \sum_{k=1}^{l} a_{j_k} \leq \sum_{k=1}^{l} \tau_k \qquad l = 1, 2, \ldots,$$

where

$$\alpha_{i_1} \geq \alpha_{i_2} \geq \dots \geq \alpha_{i_m}, \qquad a_{j_1} \geq a_{j_2} \geq \dots \geq a_{j_n},$$

and σ_k is the number of b's, and τ_k the number of β's, that are greater than or equal to k.

6. The subgraph problem. Let G be a finite directed graph, and let $e(x)$ and $i(x)$ be, respectively, the number of arcs entering and the number of arcs issuing from node x. Then the (local) degree of G at x is defined to be the pair $e(x), i(x)$.

The subgraph problem is the problem of determining conditions under which G has a subgraph H having prescribed local degrees. We consider the following generalization of this problem. Associate with each node $x \in N$ four integers $a(x)$, $b(x)$, $\alpha(x)$, $\beta(x)$, satisfying

$$0 \leq a(x) \leq b(x), \tag{17a}$$

$$0 \leq \alpha(x) \leq \beta(x), \tag{17b}$$

and determine conditions under which G has a subgraph H with local degrees $e_H(x)$, $i_H(x)$ satisfying

$$a(x) \leq e_H(x) \leq b(x), \tag{18a}$$

$$\alpha(x) \leq i_H(x) \leq \beta(x). \tag{18b}$$

To find such conditions, we convert the problem to a flow problem and apply Theorem 1. First construct from G a new directed graph G' having twice as many nodes as G but the same number of arcs: to each node x of G correspond two nodes x', x'' of G'; if (x, y) is an arc of G, then (x', y'') is an arc of G' and these are all the arcs of G'. Assign unit capacity to each arc of G'. In G', let S and T be the set of primed and double primed nodes, respectively. Next impose, for each $x' \in S$, the condition (3a) that the flow out of x' lie between $\alpha(x)$ and $\beta(x)$; similarly, for $x'' \in T$, insist that the flow into x'' lie between $a(x)$ and $b(x)$.

It is clear that an integral feasible flow f from S to T in G' singles out a subgraph H of G satisfying (18) simply by putting (x, y) in H if and only if $f(x', y'') = 1$. Conversely, of course, a subgraph H satisfying (18) produces an integral feasible flow in G'. Hence, if we let U, V be arbitrary subsets of S, T, respectively, and denote their respective complements in S, T by $\bar{U}$, $\bar{V}$, it follows from Theorem 1 and the existence of integral feasible flows that H exists if, and only if,

$$\beta(\bar{U}) + |(U, \bar{V})| \geq a(\bar{V}), \tag{19a}$$

$$b(V) + |(U, \bar{V})| \geq \alpha(U), \qquad \text{all } U \subset S,\ V \subset T, \tag{19b}$$

where $|\ |$ denotes cardinality.

Before proceeding further, let us consider inequalities (19) in the special case for which $a(x) = b(x)$, $\alpha(x) = \beta(x)$—that is, in the case for which the

local degrees of H are specified exactly. Then a necessary condition for H to exist is that $\alpha(N) = b(N)$, or, in G',

$$\alpha(S) = b(T). \tag{20}$$

On the other hand, (20) and (19b) now imply (19a), since

$$\begin{aligned}\alpha(\bar{U}) + |(U, \bar{V})| &\geq \alpha(\bar{U}) + \alpha(U) - b(V) = \alpha(S) - b(V)\\ &\geq b(T) - b(V) = b(\bar{V}),\end{aligned}$$

which is (19a) with $\alpha = \beta$, $a = b$.

Thus, (20) and (19b) are necessary and sufficient for the existence of a subgraph H having local degrees $e_H(x) = b(x)$, $i_H(x) = \alpha(x)$.

Each of the conditions (19a), (19b) is stated in terms of selections of pairs of sets. Each can, however, be simplified to a condition involving the choice of but one set. Consider (19b), for example. For given $U \subset S$, let

$$V = \{y'' \in T \mid b(y'') < |(U, y'')|\}.$$

For this pair U, V, the left-hand side of (19b) may be written as

$$\sum_{y'' \in A(U)} \min\,[b(y''), |(U, y'')|].$$

On the other hand, for fixed $U \subset S$, this sum clearly minimizes $b(V) + |(U, \bar{V})|$ over all $V \subset T$. Thus inequalities (19b) are equivalent to the inequalities

$$\sum_{y'' \in A(U)} \min\,[b(y''), |(U, y'')|] \geq \alpha(U), \qquad \text{all } U \subset S. \tag{21}$$

Similarly, (19a) reduces to

$$\sum_{y' \in B(\bar{V})} \min\,[\beta(y'), |(y', \bar{V})|] \geq a(\bar{V}), \qquad \text{all } \bar{V} \subset T. \tag{22}$$

Thus, translating (21) and (22) to conditions stated in terms of the given graph G, we have the following theorems:

THEOREM 5. *Let G be a finite directed graph with node set N, and suppose that, corresponding to each $x \in N$, there are integers $a(x)$, $b(x)$, $\alpha(x)$, $\beta(x)$ with*

$$\begin{gathered}0 \leq a(x) \leq b(x),\\ 0 \leq \alpha(x) \leq \beta(x).\end{gathered}$$

Then G has a subgraph H whose local degrees $e_H(x)$, $i_H(x)$ satisfy

$$\begin{gathered}a(x) \leq e_H(x) \leq b(x),\\ \alpha(x) \leq i_H(x) \leq \beta(x),\end{gathered}$$

if, and only if, for all $X \subset N$, we have

$$\alpha(X) \leq \sum_{y \in A(X)} \min\,[b(y), |(X, y)|], \tag{23}$$

$$a(X) \leq \sum_{y \in B(X)} \min\,[\beta(y), |(y, X)|]. \tag{24}$$

THEOREM 6 (Ore). *The finite directed graph G has a subgraph H with local degrees*

$$e_H(x) = b(x) \geq 0,$$
$$i_H(x) = \alpha(x) \geq 0,$$

if, and only if,

$$\alpha(N) = b(N) \tag{25}$$

and, for all $X \subset N$,

$$\alpha(X) \leq \sum_{y \in A(X)} \min\,[b(y), |(X, y)|]. \tag{26}$$

As a consequence of Theorem 2, we may also state the following result:

THEOREM 7. *If the finite directed graph G has subgraphs H_1, H_2, such that*

$$a(x) \leq e_{H_1}(x),\ i_{H_1}(x) \leq \beta(x),$$
$$e_{H_2}(x) \leq b(x),\ \alpha(x) \leq i_{H_2}(x),$$

where $0 \leq a(x) \leq b(x)$, $0 \leq \alpha(x) \leq \beta(x)$, then G has a subgraph H such that

$$a(x) \leq e_H(x) \leq b(x), \qquad \alpha(x) \leq i_H(x) \leq \beta(x).$$

For undirected graphs G, the (local) degree of G at x is the number of arcs incident with x, and the subgraph problem is to determine conditions under which G has a subgraph H with prescribed local degrees. In case G has only even cycles, so that the nodes of G can be partitioned into two sets S, T such that all arcs join nodes of S to those of T, the subgraph problem can be stated as a flow problem in G, and hence Theorem 1 can be applied. We know of no way, however, to make use of flow theory in the general case.

7. Systems of representatives. In our applications of the feasibility theorem thus far, the set R of intermediate nodes has been vacuous. We conclude with an application, suggested to us by Gale **(4)**, in which this will not be the case.

Let $E_1, \dots, E_n$ be subsets of a given set $E = \{e_1, \dots, e_m\}$. A list

$$D = \{e_{i_1}, \dots, e_{i_n}\}$$

of n distinct elements of E, such that $e_{i_j} \in E_j$, is a system of distinct representatives for $E_1, \dots, E_n$, in which e_{i_j} represents E_j. (A well-known theorem of P. Hall gives necessary and sufficient conditions for the existence of a system of distinct representatives.) Suppose, in addition, that $P_1, \dots, P_p$ is a partition of E, and that it is desired to establish existence conditions for a D such that the intersection of D with each P_k has cardinality between prescribed bounds. Hoffman and Kuhn **(5)** have used the duality theorem of linear-equality theory, applied to a linear-programming problem of transportation type, to prove the following theorem:

THEOREM 8 (Hoffman-Kuhn).* *Let* α_k *and* β_k, $k = 1, 2, \ldots, p$, *satisfying* $0 \leqq \alpha_k \leqq \beta_k$, *be integers associated with a partition* $P_1, \ldots, P_p$ *of a given set* $E = \{e_1, \ldots, e_m\}$. *The subsets* $E_1, \ldots, E_n$ *of* E *have a system of distinct representatives* D *satisfying* $\alpha_k \leqq |D \cdot P_k| \leqq \beta_k$, $k = 1, \ldots, p$, *if, and only if,*

$$\left|\left(\sum_{k \in U} P_k\right) \cdot \left(\sum_{j \in V} E_j\right)\right| \geqq |V| - \sum_{k \in \bar{U}} \beta_k, \tag{27}$$

$$\left|\left(\sum_{k \in U} P_k\right) \cdot \left(\sum_{j \in V} E_j\right)\right| \geqq |V| - n + \sum_{k \in U} \alpha_k, \tag{28}$$

hold for all subsets $U \subset \{1, \ldots, p\}$ *and* $V \subset \{1, \ldots, n\}$.

To establish (27) and (28) as necessary and sufficient conditions for the existence of the required system of distinct representatives, we set up the following feasibility problem. Let

$$\begin{aligned} S &= \{x_1, \ldots, x_p\}, \\ R &= \{y_1, \ldots, y_m\}, \\ T &= \{z_1, \ldots, z_n\}, \end{aligned}$$

be the nodes of a network G, and define arcs in G as follows:

$$\begin{aligned} &(x_k, y_i) \text{ is an arc if, and only if, } e_i \in P_k, \\ &(y_i, z_j) \text{ is an arc if, and only if, } e_i \in E_j. \end{aligned}$$

The capacity function is taken to be

$$\begin{aligned} c(x_k, y_i) &= 1, \\ c(y_i, z_j) &= \infty. \end{aligned}$$

With each $x_k \in S$, associate the bounds α_k, β_k on the flow leaving x_k, and similarly require that the flow into $z_j \in T$ be precisely unity ($a(z_j) = b(z_j) = 1$).

From the definition of the capacity function and the assumption that $P_1, \ldots, P_p$ is a partition of E, it follows that the amount of flow through each node $y_i \in R$ is at most one. Thus an integral feasible flow f from S to T picks out a set D fulfilling the hypotheses of the theorem:

$$D = \{e_i | f(S, y_i) = f(y_i, T) = 1\}.$$

Conversely, given a D satisfying the assumptions of the theorem, we can define an integral feasible flow f by

$$f(x_k, y_i) = \begin{cases} 1 & \text{if } e_i \in D \cdot P_k, \\ 0 & \text{otherwise;} \end{cases}$$

$$f(y_i, z_j) = \begin{cases} 1 & \text{if } e_i \text{ represents } E_j, \\ 0 & \text{otherwise.} \end{cases}$$

*It is also stated in (5) that the authors have not been able to prove this result without using the duality theorem. However, Gale has recently shown that Theorem 8 is a consequence of the circulation theorem. It is therefore not surprising that the result can be deduced from our Theorem 1.

Thus the feasibility problem in G is equivalent to the existence of a D meeting the requirements of the theorem, and we may consequently apply Theorem 1. Let $X, \bar{X}$ be a partition of the nodes of G, and set

$$S \cdot X = U, \quad R \cdot X = W, \quad T \cdot X = \bar{V},$$
$$S \cdot \bar{X} = \bar{U}, \quad R \cdot \bar{X} = \bar{W}, \quad T \cdot \bar{X} = V.$$

Then (7) and (8) become

$$\beta(\bar{U}) + c(X, \bar{X}) \geqslant |V|, \tag{29}$$
$$|V| + c(X, \bar{X}) \geqslant \alpha(U), \tag{30}$$

respectively. Since $c(y_i, z_j) = \infty$, these conditions hold automatically unless $(X, \bar{X})$ contains no arcs from R to T. Thus we may restrict attention to partitions $X, \bar{X}$ such that $B(V) \subset \bar{W}$, so that $c(X, \bar{X}) = c(U, \bar{W})$. But since the right-hand sides of (29) and (30) are independent of W, it suffices to select $\bar{W} = B(V)$. Then we have

$$c(X, \bar{X}) = c(U, B(V)) = |A(U) \cdot B(V)|.$$

Consequently, a feasible flow from S to T exists if, and only if,

$$|A(U) \cdot B(V)| \geqslant |V| - \beta(\bar{U}), \tag{31}$$
$$|A(U) \cdot B(V)| \geqslant \alpha(U) - |\bar{V}|, \quad \text{all } U \subset S, \quad V \subset T. \tag{32}$$

Replacing $|\bar{V}|$ by $n - |V|$ in (32) and translating (31) and (32) into set-theoretic statements yield (27) and (28), respectively. Thus (27) is a necessary and sufficient condition that there be a system of distinct representatives D such that $|D \cdot P_k| \leqslant \beta_k$, whereas (28) is a necessary and sufficient condition that there be a system of distinct representatives D with $|D \cdot P_k| \geqslant \alpha_k$.

References

1. L. R. Ford, Jr. and D. R. Fulkerson, *Maximal flow through a network*, Can. J. Math., *8* (1956), 399–404.
2. ——— *A simple algorithm for finding maximal network flows and an application to the Hitchcock problem*, Can. J. Math., *9* (1957), 210–18.
3. ——— *Network flow and systems of representatives*, Can. J. Math., *10* (1958), 78–85.
4. D. Gale, *A theorem on flows in networks*, Pac. J. Math., *7* (1957), 1073–82.
5. A. J. Hoffman and H. W. Kuhn, *On systems of distinct representatives*, in H. W. Kuhn and A. W. Tucker (eds.), *Linear inequalities and related systems*, Annals of Mathematics Study No. 38, (Princeton, 1956).
6. O. Ore, *Studies on directed graphs, I*, Ann. Math., *63* (1956), 383–406.
7. ——— *Studies on directed graphs, II*, Ann. Math., *64* (1956), 142–53.
8. ——— *Graphs and subgraphs*, Trans. Amer. Math. Soc., *84* (1957), 109–37.
9. H. J. Ryser, *Combinatorial properties of matrices of zeros and ones*, Can. J. Math., *9* (1957), 371–7.
10. W. T. Tutte, *The factorization of linear graphs*, J. Lond. Math. Soc., *22* (1947), 107–11.
11. ——— *The 1-factors of oriented graphs*, Proc. Amer. Math. Soc., *4* (1953), 922–30.
12. ——— *The factors of graphs*, Can. J. Math., *4* (1952), 314–29.

The Rand Corporation

THE REGULAR MAPS ON A SURFACE OF GENUS THREE

F. A. SHERK

Introduction. A considerable volume of research on the theory of regular maps is now in existence. Systematic enumerations of regular maps on the surfaces of genus 1 and 2 were begun by Brahana (**1**; **2**) and completed by Coxeter (**6**; **7**, p. 141). In addition Coxeter enumerated the regular maps on the simplest non-orientable surfaces (**7**, pp. 116, 139), and constructed tables of some interesting families of regular maps (**3**; **7**, p. 140).

Most of the regular maps on a surface of genus 3 have appeared in these papers, but no systematic enumeration of them seems to have been attempted. The ultimate goal of this paper is a complete list of these regular maps. However, the families of maps $\{j\cdot p, q\}$ and $\{j\cdot p, j\cdot q\}$ which are defined in § 4 and listed in Tables I and II are of considerable interest in themselves. Also of some importance is the complete list of regular maps of type $\{p, 3\}$ with six or fewer faces (§ 5 and Table III).

A method of deriving regular maps by identification of faces in a regular tessellation is introduced in § 2 and used in §§ 5 and 7. Although cumbersome in some cases, it is the only reliable tool which has yet been developed for completing a list of regular maps of genus $p > 1$ (Brahana's method (**2**, pp. 281–4) is dependent upon the completeness and accuracy of permutation group tables).

1. Elementary concepts and results. A *map* is a partitioning of an unbounded surface into N_2 simply-connected, non-overlapping regions called *faces* by means of N_1 lines called *edges*. The N_0 intersections of the edges are called *vertices*.

The Euler-Poincaré characteristic

1.1 $$\chi = N_0 - N_1 + N_2$$

has the same value for every map drawn on this surface. If the surface is orientable, then

1.2 $$\chi = 2 - 2\mathrm{p},$$

where p is the genus of the surface.

Received July 25, 1958. The bulk of these results form a part of the author's Ph.D. thesis, which was written under the supervision of Professor H. S. M. Coxeter. The author wishes to express his sincere thanks to Professor Coxeter for his kind and generous help.

The research was carried on while the author held a research studentship from the National Research Council of Canada.

To every map there corresponds a *dual* map having N_0 faces, one surrounding each vertex or the original map, N_1 edges, one crossing each edge of the original map, and N_2 vertices, one contained in the interior of each face of the original map (**5**, p. 6).

With any map there is associated a group of transformations which leave the map invariant and preserve incidences, that is, a group of *automorphisms* (**7**, p. 100). An automorphism is determined by its effect on any one face. Suppose that the group contains, in particular, two automorphisms R and S, the first of which cyclically permutes the edges bounding a face F, while the other cyclically permutes the edges which meet at a vertex V of F. A map containing these two automorphisms is said to be *regular*.

It is immediately evident that if the face F is p-sided, and if q edges meet at the vertex V, then every face of the regular map is p-sided and exactly q edges meet at every vertex. Thus the regular map is composed of p-gons, q meeting at each vertex. Such a map is said to be a "map of type $\{p, q\}$," in analogy with Schläfli's notation for a regular polyhedron (**5**, p. 14). The dual map is of type $\{q, p\}$ and is, of course, also regular. It also follows from the definition of a regular map that the group of the map is transitive on its vertices, edges, and faces.

Suppose that we divide the surface of the regular map of type $\{p, q\}$ into pN_2 triangles by adding to the map the lines which join the vertices of each face to the corresponding vertex of the dual map (cf. Figure 1 for the case of a map of type $\{6, 3\}$). Thus each face of the map is made up of p triangles,

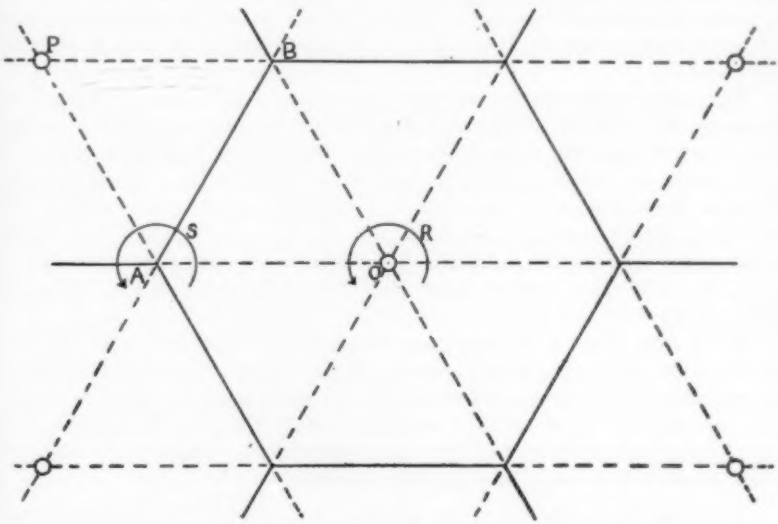

FIGURE 1

each edge borders on 2 triangles, and each vertex is surrounded by $2q$ triangles. It follows that

1.3 $$pN_2 = 2N_1 = qN_0.$$

Accordingly if the map has N_1 edges it has $2N_1/p$ faces and $2N_1/q$ vertices. Substituting in formula 1.1, we have for the surface of the regular map:

1.4 $$\chi = 2N_1\left(\frac{1}{p} + \frac{1}{q} - \frac{1}{2}\right).$$

We define the *group* of a regular map to be the group which is generated by the automorphisms R and S. Examining Figure 1, we note that the automorphism* RS interchanges the triangles OAB and PAB. Thus RS is of period 2. It is easy to see that this result is true for any regular map; the group of a regular map of type $\{p, q\}$ must satisfy the relations

1.5 $$R^p = S^q = (RS)^2 = E,$$

where E denotes the identity element. These relations are sufficient to define the group if the surface is simply-connected, but in any other case at least one extra relation is needed.

Looking again at Figure 1, we note that the edge AB is carried into itself by two automorphisms in the group, namely E and RS. When the surface on which the map lies is non-orientable, the group contains two other automorphisms which carry AB into itself. One of these will leave A and B invariant, interchanging O and P, while the other leaves O and P invariant and interchanges A and B. These automorphisms are called *reflections* since they operate in a manner analogous to the reflections of the Euclidean plane (**5**, p. 75). Since the group is transitive on the edges of the map it must be of order $4N_1$.

Any regular map whose automorphisms include reflections is said to be *reflexible* (**7**, p. 101). Certain non-reflexible regular maps do exist. Coxeter (**6**, p. 26; **7**, pp. 103, 107) exhibited the non-reflexible regular maps on a surface of genus 1 and stated that no others were known (**7**, p. 102). However, Frucht (**9**) discovered a non-reflexible regular map on a surface of genus 55 which is the embedding in that surface of a one-regular graph of degree three. Any non-reflexible regular map must lie on an orientable surface, since the group of a regular map on a non-orientable surface must contain reflections (**7**, p. 101).

If the map is on a non-orientable surface, or if it is non-reflexible, the group of the map is the complete group of automorphisms. Every map which is reflexible and lies on an orientable surface has a larger group of automorphisms which we shall call the *extended group* of the map (**4**, p. 125). The extended group includes reflections and is therefore of order $4N_1$. It contains "the group of the map" as a subgroup of index 2.

*By RS, the product of R and S, we mean the automorphism which is achieved by performing R first and then performing S.

The automorphisms that comprise the group of an orientable regular map are called *rotations*. By 1.3 the order of the group may be expressed in the forms pN_2 or qN_0 as well as in the form $2N_1$.

In virtue of relations 1.4 and 1.2, any regular map of type $\{p, q\}$ which has N_1 edges and is on an orientable surface is on a surface of genus

$$1.6 \qquad \mathrm{p} = 1 - N_1\left(\frac{1}{p} + \frac{1}{q} - \frac{1}{2}\right).$$

The expressions "regular map on a surface of genus p" will now be shortened to "regular map of genus p."

The regular maps of genus zero are simply the projections on concentric spheres of the 5 convex regular polyhedra, $\{3, 3\}$, $\{4, 3\}$, $\{3, 4\}$, $\{5, 3\}$, and $\{3, 5\}$. together with the "dihedral" maps $\{p, 2\}$ ($p > 1$) and their duals $\{2, p\}$. The groups of the regular maps of genus zero are the well-known polyhedral rotation groups (**11**, pp. 10–20; **5**, pp. 45–7), denoted by the symbols $[p, q]^+$ (**7**, p. 38), whose abstract definitions are given by 1.5 with appropriate values for p and q. Thus in the case of the regular maps of genus zero, the relations 1.5 are sufficient, as well as necessary, to define the group.

2. The regular tessellations. The above description of a regular map can be extended to include regular maps on an infinite surface. Thus, for example, we have in the Euclidean plane the regular maps $\{4, 4\}$, $\{6, 3\}$, and $\{3, 6\}$, more commonly called *regular tessellations* (**5**, pp. 58, 59). There are also regular tessellations in the hyperbolic plane (**7**, p. 53); they are of type $\{p, q\}$ for all p and q such that $(\mathrm{p} - 2)\ (q - 2) > 4$. The regular tessellations on the sphere are just the regular maps of genus zero. All regular tessellations are simply-connected maps.

As in the case of the regular maps on a sphere, the relations 1.5 are sufficient to define the group of a regular tessellation. It follows that the group of a regular map of type $\{p, q\}$ on an orientable surface is a factor group of the group of the regular tessellation $\{p, q\}$. It is also true that the plane of the tessellation is a universal covering surface for the surface in question (**7**, pp. 25, 26). These facts suggest a method of discovering regular maps. Beginning with a regular tessellation $\{p, q\}$ we add further relations to those of 1.5 by abstractly identifying certain faces of the tessellation (the exact procedure in this step will be outlined in the proof of Theorem 3). If the added relations do not effect the periods of R, S, and RS, and if they are sufficient to make the resulting group finite, let us say of order g, a regular map of type $\{p, q\}$ has been discovered. It has g/q vertices, $g/2$ edges, and g/p faces. It lies on a surface of genus

$$1 - g/2\left(\frac{1}{p} + \frac{1}{q} - \frac{1}{2}\right)$$

(cf. 1.6).

Furthermore, the above method will establish the existence or non-existence of all regular maps of type $\{p, q\}$ with given group order.

3. Some general lemmas. The following lemmas form the essential groundwork for all our results.

LEMMA 1. *For any map of type $\{p, q\}$ on a surface of Euler-Poincaré characteristic $\chi < 1$, min $(p, q) \geqslant 3$.*

Proof. Consider a map of type $\{p, q\}$ which has k faces. It follows from 1.3 and 1.1 that

$$\chi = \frac{pk}{q} - \frac{pk}{2} + k.$$

Rearranging this equation, we obtain

$$k - \chi = pk\,(q - 2)/2q.$$

If $q \leqslant 2$, then $k - \chi \leqslant 0$ and $\chi \geqslant k \geqslant 1$. Thus if $\chi < 1$, $q \geqslant 3$. A similar argument holds for p when one considers the dual map, of type $\{q, p\}$.

LEMMA 2. *If two edges belonging to the same face of a regular map are identified, the map has only one face.*

Proof. We noted earlier that the group of a regular map is transitive on the edges of that map. Thus if two edges of a face are identified, then all the other edges of that face are also identified in pairs; the result is a one-faced map.

LEMMA 3. *If exactly two distinct faces come together at a vertex of a regular map of type $\{p, q\}$, the map is 2-faced, q is even, and the faces alternate around the vertex.*

Proof. If a face is contiguous to itself around a vertex, then by Lemma 2 the map is one-faced, contrary to our hypothesis. Thus q is even and the faces, α and β say, which surround a vertex alternate around that vertex (cf. Figure 2, where α and β surround the vertex V). Now consider any edge VV' (Figure 2). This edge borders on α and β, and hence α and β alternate around V' as well as around V. This happens at every vertex since the group of the map is transitive on its edges. Therefore α and β are the only faces.

LEMMA 4. *A one-faced map of type $\{p, q\}$ is regular if, and only if, one of the following two conditions is satisfied:*

(i) *$\frac{1}{2}p$ is an even integer and $q = p$;*
(ii) *$\frac{1}{2}p$ is an odd integer and $q = \frac{1}{2}p$.*

Proof. The single face of a one-faced map must have an even number of edges since these edges are identified in pairs to form the edges of the map. Thus $p = 2n$, where n is some integer, and the group of the one-faced regular map $\{2n, q\}$ is the cyclic group of order $2n$ generated by the rotation R of

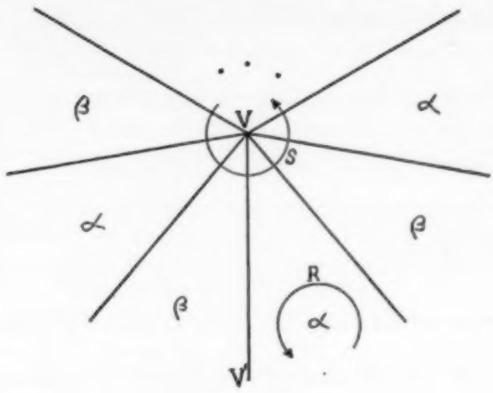

FIGURE 2

§ 1. Now the group of any regular map may be expressed in terms of the generators R and $T = RS$ instead of the R and S used earlier (**2**, p.269). The three relations of 1.5 are then equivalent to

3.1 $$R^p = T^2 = (RT)^q = E.$$

In the present case, T must be expressible in terms of R, and since $T^2 = E$, $T = R^n$. The existence of a regular map of type $\{2n, q\}$ depends upon the period of RT, which must be q. But $RT = R^{n+1}$. Hence if the map is regular

$$(RT)^q = R^{q(n+1)} = E = R^{2n},$$

and therefore $2n|q(n + 1)$. Now $(n, n + 1) = 1$, so that if n is even, $2n \mid q$, while if n is odd, $n \mid q$. Since the map has only n edges, $q \leq 2n$. Thus if n is even, $q = 2n$, while if n is odd,

$$(RT)^n = R^{(n+1)n} = (R^{2n})^{\frac{1}{2}(n+1)} = E,$$

and thus $q \mid n$. But $n \mid q$, therefore $q = n$. Conversely, any one-faced map of type {4p, 4p} or {4p + 2, 2p + 1} (p = 0, 1, 2, . . .) is regular.

It is easily seen from 1.6 that the above two one-faced regular maps lie on a surface of genus p.

LEMMA 5. *If the rotation $R^h (1 < h < p$, where p is the period of $R)$ carries a vertex, edge, or face of a regular map into itself, while any rotation $R^i(0 < i < h)$ does not do so, then $p \equiv 0 \pmod{h}$.*

Proof. The integer p may be put into the form

$$p = mh + n$$

where m and n are integers and $0 \leq n < h$. Since both R^h and R^p $(= E)$ carry the vertex, edge, or face into itself, so also must R^n. Therefore $n = 0$.

LEMMA 6. *The abstract definition**

3.2 $R^{jp} = S^q = (RS)^2 = E, \quad R^p \rightleftarrows S \quad (0 < (p-2)(q-2) < 4)$

is significant only if $j \mid Q$, *where* $Q = 4q/[4-(p-2)(q-2)]$, *and then defines a group of order* jpQ.

Proof. If we exclude the first relation in 3.2 we have

3.3 $$S^q = (RS)^2 = E, \quad R^p \rightleftarrows S.$$

If $SR = T$, this becomes

$$T^2 = S^q = E, \qquad (TS)^p = (ST)^p.$$

These relations define the group $\langle\langle 2, q \mid p \rangle\rangle$ of order PQ^2, which was introduced by Coxeter and Moser (**7**, p. 79). The period of R $(= TS^{-1})$ is pQ (**7**, p. 71) and therefore the abstract definition 3.2 is significant only if this period is a multiple of jp. If we add to 3.3 the relation $R^{jp} = E$, where $j \mid Q$, the only effect is to change the period of R to jp; the periods of S and RS will remain unchanged. Now it is easily shown that the number of cosets of $\{R\}$ in 3.2 remains the same, no matter what the particular choice of j is. When $j = 1$, the group is $[p, q]^+$, the group of the regular map $\{p, q\}$ and $\{R\}$ has Q cosets (**7**, p. 38). Thus the group defined by 3.2 has order jpQ.

To the group defined by 3.2 or

$$(TS)^p = (ST)^p = Z, \qquad T^2 = S^q = Z^j = E,$$

we assign the symbol $\langle\langle 2, q \mid p; j \rangle\rangle$. In particular, $\langle\langle 2, q \mid p; 1 \rangle\rangle = [p, q]^+$.

4. Two new families of regular maps. Coxeter and Moser (**7**, § 8.8) introduced the regular map $\{p+p, q\}$ $(0 < (p-2)(q-2) < 4)$ and its dual $\{q, p+p\}$, whose group has the abstract definition.

$$R^{2p} = T^2 = (RT)^q = (R^pT)^2 = E,$$

or, in terms of R and $S = RT$,

$$R^{2p} = S^q = (RS)^2 = E, \qquad R^p \rightleftarrows S.$$

We generalize this notion by considering the regular map of type $\{jp, q\}$ $(0 < (p-2)(q-2) < 4)$ and its dual, of type $\{q, jp\}$, whose group G has the following property: the centre of G is a cyclic group, generated by R^p, where R has its usual meaning as a generator of the group. Such a group G will satisfy the following four relations:

4.1 $$R^{jp} = S^q = (RS)^2 = E, \qquad R^p \rightleftarrows S.$$

By Lemma 6, these relations are significant only if $j \mid Q$, where $Q = 4q/[4 - (p-2)(q-2)$, and then they define the group $\langle\langle 2, q \mid p; j \rangle\rangle$ of order jpQ.

*The notation $A \rightleftarrows B$ means that A and B commute.

Now the central quotient group of G is the group $[p, q]^+$ of order pQ. Thus G is of order jpQ, which is precisely the order of the group $\langle\langle 2, q \mid p; j\rangle\rangle$. Therefore the relations 4.1 define G.

To the above map of type $\{jp, q\}$ we assign the symbol $\{j\cdot p, q\}$, and denote its dual by $\{q, j\cdot p\}$. The map occurs for all integer values $p \geqslant 2$, $q \geqslant 2$ and $j > 0$ satisfying the two conditions $0 \leqslant (p - 2)(q - 2) < 4$ and $j \mid Q$. Its group is $\langle\langle 2, q \mid p; j\rangle\rangle$.

In particular, the maps $\{2\cdot p, q\}$ and $\{q, 2\cdot p\}$ are the $\{p + p, q\}$ and $\{q, p + p\}$ respectively of Coxeter and Moser (**7**, § 8.8) who pointed out that these maps may be drawn on a two-sheeted Riemann surface of the proper genus in a remarkably symmetrical manner. This construction is capable of generalization to the case of the regular maps $\{j.p, q\}$ and $\{q, j.p\}$ $(j \neq 2)$.

In the proof of Lemma 6 it was shown that when $j = Q$, the groups 3.2 are the groups $\langle\langle 2, q \mid p\rangle\rangle$ of order pQ^2. They were shown by Coxeter and Moser (**7**, pp. 79–80) to be the groups of the regular complex polygons $2\{2p\}q$, discovered by Shephard (**12**, p. 92). When these complex polygons are compared with the corresponding regular maps $\{pQ\cdot q, p\}$, it can be shown by the proper interpretation of the group generators in each case that the vertices and edges of the polygon form the same graph as the vertices and edges of the map. Thus the map may be regarded as a real representation of the complex polygon.

Generalizing in another direction, we consider the regular map of type $\{jp, jq\}$ $(0 \leqslant (p - 2)(q - 2) < 4)$ and its dual, of type $\{jq, jp\}$, whose group G' has the following properties: the centre of G' is a cyclic group generated by R^p; and $R^p = S^q$, where R and S have their usual meanings as generators of G'. Such a group will have among its defining relations the following:

$$R^p = S^q = Z, \qquad (RS)^2 = Z^j = E. \tag{4.2}$$

These relations are significant only if

$$j \,\Big|\, \frac{p+q}{q}\, Q,$$

where $Q = 4q/[4 - (p - 2)(q - 2)]$, and then they define the group $\langle p, q \mid 2; j\rangle$ of order jpQ (**7**, pp. 71–3). This is a factor group of Miller's group $\langle p, q \mid 2\rangle$ which is defined by the relations

$$R^p = S^q, \qquad (RS)^2 = E.$$

Now the centre $\{R^p\}$ of G' is of order j, and the central quotient group of G' is $[p, q]^+$, of order pQ. Thus G' is of order jpQ, which is precisely the order of the group $\langle p, q \mid 2; j\rangle$. Therefore G' is defined by 4.2.

To the above type of map whose group is $\langle p, q \mid 2; j\rangle$ we assign the symbol $\{j\cdot p, j\cdot q\}$, and denote its dual by $\{j\cdot q, j\cdot p\}$. The map occurs for all integer

values $p < 2$, $q > 2$ and $j > 0$ satisfying the two conditions $0 < (p-2)(q-2) < 4$ and

$$j \left| \frac{p+q}{q} Q. \right.$$

The members $\{(r+1)\cdot(r-1),\ (r+1)\cdot 2\}$ of this family were noted by Coxeter and Moser (**7**, p. 114).

The regular maps $\{p, q\}$ of genus zero are members of the family $\{j\cdot p, q\}$ as well as of the family $\{j\cdot p, j\cdot q\}$. With these exceptions, all the regular maps $\{j\cdot p, q\}$ and $\{j\cdot p, j\cdot q\}$ $(p > q)$ are listed in Tables I and II respectively. The sixth column in both tables exhibits some interesting isomorphisms between the group of the map and certain well-known groups. The information for the sixth column of Table I was kindly supplied by W. O. J. Moser; in the case of Table II the source is 7, § 6.6.

TABLE I

THE REGULAR MAPS $\{j\cdot p, q\}$ $(j > 2)$

Map	N_0	N_1	N_2	Genus	Group	Order
$\{j\cdot 2, q\}$ $(j\vert q)$	$2j$	jq	q	$\frac{1}{2}(j-1)(q-2)$	$\langle\langle 2, q \mid 2; j\rangle\rangle$	$2jq$
$\{q\cdot 2, q\}$	$2q$	q^2	q	$\frac{1}{2}(q-1)(q-2)$	$\langle\langle 2, q \mid 2\rangle\rangle$	$2q^2$
$\{2\cdot p, 2\}$	$2p$	$2p$	2	0	$\langle\langle 2, 2 \mid p\rangle\rangle \cong \mathfrak{D}_{2p}$	$4p$
$\{2\cdot 3, 3\}$	8	12	4	1	$\langle\langle 2, 3 \mid 3; 2\rangle\rangle \cong \mathfrak{A}_4\times\mathfrak{C}_2$	24
$\{4\cdot 3, 3\}$	16	24	4	3	$\langle\langle 2, 3 \mid 3\rangle\rangle$	48
$\{2\cdot 4, 3\}$	16	24	6	2	$\langle\langle 2, 3 \mid 4; 2\rangle\rangle$	48
$\{3\cdot 4, 3\}$	24	36	6	4	$\langle\langle 2, 3 \mid 4; 3\rangle\rangle$	72
$\{6\cdot 4, 3\}$	48	72	6	10	$\langle\langle 2, 3 \mid 4\rangle\rangle$	144
$\{2\cdot 5, 3\}$	40	60	12	5	$\langle\langle 2, 3 \mid 5; 2\rangle\rangle \cong \mathfrak{A}_5\times\mathfrak{C}_2$	120
$\{3\cdot 5, 3\}$	60	90	12	10	$\langle\langle 2, 3 \mid 5; 3\rangle\rangle \cong \mathfrak{A}_5\times\mathfrak{C}_3$	180
$\{4\cdot 5, 3\}$	80	120	12	15	$\langle\langle 2, 3 \mid 5; 4\rangle\rangle$	240
$\{6\cdot 5, 3\}$	120	180	12	25	$\langle\langle 2, 3 \mid 5; 6\rangle\rangle \cong \mathfrak{A}_5\times\mathfrak{C}_6$	360
$\{12\cdot 5, 3\}$	240	360	12	55	$\langle\langle 2, 3 \mid 5\rangle\rangle$	720
$\{2\cdot 3, 4\}$	12	24	8	3	$\langle\langle 2, 4 \mid 3; 2\rangle\rangle \cong \mathfrak{S}_4\times\mathfrak{C}_2$	48
$\{4\cdot 3, 4\}$	24	48	8	9	$\langle\langle 2, 4 \mid 3; 4\rangle\rangle \cong \mathfrak{S}_4\times\mathfrak{C}_4$	96
$\{8\cdot 3, 4\}$	48	96	8	21	$\langle\langle 2, 4 \mid 3\rangle\rangle$	192
$\{2\cdot 3, 5\}$	24	60	20	9	$\langle\langle 2, 5 \mid 3; 2\rangle\rangle \cong \mathfrak{A}_5\times\mathfrak{C}_2$	120
$\{4\cdot 3, 5\}$	48	120	20	27	$\langle\langle 2, 5 \mid 3; 4\rangle\rangle$	240
$\{5\cdot 3, 5\}$	60	150	20	36	$\langle\langle 2, 5 \mid 3; 5\rangle\rangle \cong \mathfrak{A}_5\times\mathfrak{C}_5$	300
$\{10\cdot 3, 5\}$	120	300	20	81	$\langle\langle 2, 5 \mid 3; 10\rangle\rangle \cong \mathfrak{A}_5\times\mathfrak{C}_{10}$	600
$\{20\cdot 3, 5\}$	240	600	20	171	$\langle\langle 2, 5 \mid 3\rangle\rangle$	1200

5. Regular maps of type $\{p, 3\}$. In some respects the most interesting regular maps are those which have 3 faces at a vertex. We shall now proceed to enumerate these maps when the number of faces is small.

To facilitate reference to it, a map of type $\{p, q\}$ having k faces will be denoted by the symbol ${}^k\{p, q\}$. In particular we shall now study the regular maps ${}^k\{p, 3\}$ for small values of k.

TABLE II

THE REGULAR MAPS $\{j\cdot p, j\cdot q\}$ $(p > q; j > 2)$

Map	N_0	N_1	N_2	Genus	Group	Order
$\{j\cdot p, j\cdot 2\}$ $(j \mid p+2)$	p	jp	2	$\frac{1}{2}(j-1)p$	$\langle p, 2 \mid 2; j\rangle$	$2jp$
$\{(p+2)\cdot p, (p+2)\cdot 2\}$	p	$p(p+2)$	2	$\frac{1}{2}p(p+1)$	$\langle p, 2 \mid 2\rangle$	$2p(p+2)$
$\{2\cdot 3, 2\cdot 3\}$	4	12	4	3	$\langle 3, 3 \mid 2; 2\rangle \cong \mathfrak{A}_4\times\mathfrak{C}_2$	24
$\{4\cdot 3, 4\cdot 3\}$	4	24	4	9	$\langle 3, 3 \mid 2; 4\rangle \cong \mathfrak{A}_4\times\mathfrak{C}_4$	48
$\{8\cdot 3, 8\cdot 3\}$	4	48	4	21	$\langle 3, 3 \mid 2\rangle$	96
$\{2\cdot 4, 2\cdot 3\}$	8	24	6	6	$\langle 4, 3 \mid 2; 2\rangle$	48
$\{7\cdot 4, 7\cdot 3\}$	8	84	6	36	$\langle 4, 3 \mid 2; 7\rangle$	168
$\{14\cdot 4, 14\cdot 3\}$	8	168	6	78	$\langle 4, 3 \mid 2\rangle$	336
$\{2\cdot 5, 2\cdot 3\}$	20	60	12	15	$\langle 5, 3 \mid 2; 2\rangle \cong \mathfrak{A}_5\times\mathfrak{C}_2$	120
$\{4\cdot 5, 4\cdot 3\}$	20	120	12	45	$\langle 5, 3 \mid 2; 4\rangle \cong \mathfrak{A}_5\times\mathfrak{C}_4$	240
$\{8\cdot 5, 8\cdot 3\}$	20	240	12	105	$\langle 5, 3 \mid 2; 8\rangle \cong \mathfrak{A}_5\times\mathfrak{C}_8$	480
$\{16\cdot 5, 16\cdot 3\}$	20	480	12	225	$\langle 5, 3 \mid 2; 16\rangle \cong \mathfrak{A}_5\times\mathfrak{C}_{16}$	960
$\{32\cdot 5, 32\cdot 3\}$	20	960	12	465	$\langle 5, 3 \mid 2\rangle$	1920

From Lemma 4 we deduce

THEOREM 1. *The only regular map* $^1\{p, 3\}$ *is the map* $\{6, 3\}_{1,0}$ *of genus* 1 (**6**, p. 25).

Lemmas 2 and 3 imply

THEOREM 2. *There is no regular map* $^2\{p, 3\}$.

Turning now to the case $k = 3$, we exhibit in Figure 3 a part of the regular tessellation $\{p, 3\}$. In virtue of Lemmas 2 and 3 the three faces of a regular

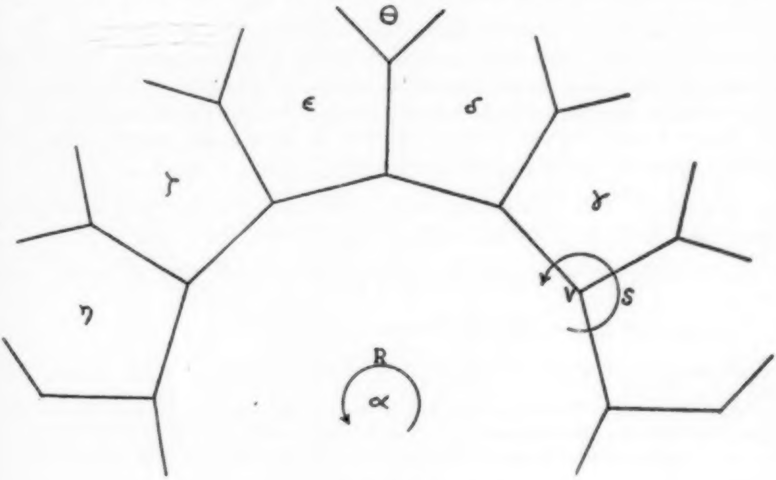

FIGURE 3

map ${}^3\{p, 3\}$ are situated in the manner of the faces α, β, and γ. Since the map is 3-faced, δ must be identified with β. Thus, representing faces by right cosets* of $\{R\}$, where the automorphisms R and S act in the indicated manner, we have

$$\{R\}SR^2 = \{R\}S.$$

In particular, there is an integer l such that

$$SR^2 = R^lS.$$

Thus the generators of the group of ${}^3\{p, 3\}$ must satisfy this relation as well as

$$R^p = S^3 = (RS)^2 = E.$$

It follows that

$$\begin{aligned} R^l &= SR^2S^{-1} = S^{-2}R^2S^2 = S^{-1}RSR^2S^{-1}R^{-1}S \\ &= S^{-1}RR^lR^{-1}S = S^{-1}R^lS = R^2, \end{aligned}$$

and the extra relation reduces to

$$R^2 \rightleftarrows S.$$

Moreover, Lemma 5 shows that p is even. Thus the abstract definition

5.1 $$R^p = S^3 = (RS)^2 = E, \qquad R^2 \rightleftarrows S$$

is a special case of 3.2, and Lemma 6 shows that $p = 2$ or 6. Thus we have

THEOREM 3. *There are exactly two regular maps* ${}^3\{p, 3\}$, *namely* $\{2, 3\}$ *of genus zero and* $\{3\cdot 2, 3\}$ *of genus* 1.

In the notation of Coxeter (**6**, p. 25), $\{3\cdot 2, 3\}$ is the map $\{6, 3\}_{1,1}$.

Lemma 6 also shows that the identification of faces carried out in the above case can yield only a 3-faced regular map (of type $\{p, 3\}$). Thus the four faces of any map ${}^4\{p, 3\}$ are situated in the manner of α, β, γ, and δ in Figure 3, and ϵ must be identical with β. By similar reasoning to that used in proving Theorem 3, we now have

THEOREM 4. *There are exactly three regular maps* ${}^4\{p, 3\}$, *namely* $\{3, 3\}$ *of genus zero,* $\{2\cdot 3, 3\}$ *of genus* 1, *and* $\{4\cdot 3, 3\}$ *of genus* 3.

In the notation of Coxeter (**6**, p. 25; cf. **7**, p. 116), $\{2\cdot 3, 3\}$ is the map $\{6, 3\}_{2,0}$.

Turning now to the case ${}^5\{p, 3\}$, we prove

THEOREM 5. *There is no regular map* ${}^5\{p, 3\}$.

*Since the rotation R carries α into itself, α may be represented in the group by the subgroup $\{R\}$ while the other faces are represented by right cosets of $\{R\}$ (**2**, p. 270). Thus there is a (1, 1) correspondence between the faces of a regular map and the right cosets of $\{R\}$ in its group.

The reader is requested to insert the letter β in the face to the right of α (Figure 3).

Proof. Suppose that a regular map ${}^5\{p, 3\}$ exists. Then in view of the results of the two previous theorems, its faces must be situated in the manner of α, β, γ, δ, and ϵ in Figure 3, and ζ must be identical with β. The group of the map must therefore satisfy the relations

$$5.2 \qquad R^p = S^3 = (RS)^2 = E, \qquad R^4 \rightleftarrows S,$$

where $p \equiv 0 \pmod 4$. But by Lemma 6, 5.2 defines a group of order $6p$, while a regular map ${}^5\{p, 3\}$ must have a group of order $5p$. The group defined by 5.2 cannot have a factor group of order $5p$; hence there is no regular map ${}^5\{p, 3\}$.

It was noted in the above proof that 5.2 defines a group of order $6p$. Hence the identification of ζ with β in Figure 3 yields regular maps ${}^6\{p, 3\}$. We ask if any other identification of faces in the tessellation $\{p, 3\}$ will yield 6-faced regular maps. The only other possible arrangement is to let α, β, γ, δ, ϵ, and ζ be the 6 faces and identify η with β. This gives rise to a group satisfying the relations

$$5.3 \qquad R^p = S^3 = (RS)^2 = E, \qquad R^5 \rightleftarrows S$$

where $p \equiv 0 \pmod 5$. By Lemma 6 this defines a group of order $12p$, while a regular map ${}^6\{p, 3\}$ must have a group of order $6p$. Thus relations 5.3 are insufficient to define the group which we seek, and we must add a further relation. Since the regular map we seek is 6-faced, the face θ of Figure 3 must be identified with α, β, γ, δ, ϵ, or ζ. Each face is surrounded by 5 different faces, and therefore θ can only be identified with β; in symbols

$$\{R\}SR^{-1}SR^2 = \{R\}S.$$

In particular, there is an integer l such that $SR^{-1}SR^2 = R^lS$. However, if we add this relation to those of 5.3 and enumerate right cosets of $\{R\}$ by the Todd-Coxeter method (**7**, p. 12), a collapse occurs in the tables which reduces the number of cosets of $\{R\}$ to one. Thus we eliminate the possibility of a group of order $6p$.

Since the Todd-Coxeter method will be employed many times in similar situations, it is perhaps advisable to exhibit the tables in this case. They are

$R\,R\,R\,R\,R \ldots R$	$S\,S\,S$	$R\,S\,R\,S$	$R^5\,S$	$S\,R^5$
1 1 1 1 1 1 1 1	1 2 3 1	1 1 2 3 1	1 1 2	1 2 2
2 3 4 5 6 2 6 2	4 6 5 4	3 4 6 2 3	2 2 3	2 3 3
		4 5 4 5 4	3 3 1	3 1 1
		5 6 5 6 5		

$S\,R^{-1}\,S\,R^2$	$R^l\,S$
1 2 6 5 2	1 1 2
2 3 2 3 5	2 6 5
3 1 1 2 4	3 5 4

The table for the fifth relation indicates that R^l carries coset 2 into coset 6

and at the same time carries coset 3 into coset 5. Transferring this information to the table for the first relation, we see that cosets 2, 3, 4, 5, and 6 are identical. But the table for the second relation then indicates that coset 2 = coset 1 and the collapse is complete. It is important to notice that the enumeration of cosets is carried out without knowing the specific values of p and l. This time-saving fact should be kept in mind and applied to any particular case when an enumeration of cosets is desired in the following pages.

Collecting the above results, and taking Lemma 6 into account, we have

THEOREM 6. *The only regular maps* ${}^6\{p, 3\}$ *are* $\{4, 3\}$ *of genus zero,* $\{2\cdot 4, 3\}$ *of genus* 2, $\{3\cdot 4, 3\}$ *of genus* 4, *and* $\{6\cdot 4, 3\}$ *of genus* 10.

It is not difficult to classify completely the regular maps ${}^k\{p, 3\}$ for other small values of k by using the above methods. For example, it can be shown quite easily that there is only one regular map ${}^7\{p, 3\}$, namely (**6**, p. 25) the map $\{6, 3\}_{2,1}$ of genus 1. For the present, however, we shall confine ourselves to the following result, obtained by an examination of the proofs of Theorems 3–6.

THEOREM 7. *The faces of any regular map,* ${}^k\{p, 3\}$ $(k > 6)$ *are surrounded by at least five other distinct faces.*

The regular maps ${}^k\{p, 3\}$ $(k \leqslant 6)$ are listed in Table III.

TABLE III

THE REGULAR MAPS OF TYPE $\{p, 3\}$ WITH SIX OR FEWER FACES

Symbol	N_0	N_1	N_2	Genus	Group
$\{6, 3\}_{1,0}$	2	3	1	1	$\mathfrak{C}_6$
$\{2, 3\}$	2	3	3	0	$[2, 3]^+ \cong \mathfrak{D}_3$
$\{3\cdot 2, 3\}$	6	9	3	1	$\langle\langle 2, 3 \mid 2\rangle\rangle$
$\{3, 3\}$	4	6	4	0	$[3, 3]^+ \cong \mathfrak{A}_4$
$\{2\cdot 3, 3\}$	8	12	4	1	$\langle\langle 2, 3 \mid 3; 2\rangle\rangle \cong \mathfrak{A}_4 \times \mathfrak{C}_2$
$\{4\cdot 3, 3\}$	16	24	4	3	$\langle\langle 2, 3 \mid 3\rangle\rangle$
$\{4, 3\}$	8	12	6	0	$[4, 3]^+ \cong \mathfrak{S}_4$
$\{2\cdot 4, 3\}$	16	24	6	2	$\langle\langle 2, 3 \mid 4; 2\rangle\rangle$
$\{3\cdot 4, 3\}$	24	36	6	4	$\langle\langle 2, 3 \mid 4; 3\rangle\rangle$
$\{6\cdot 4, 3\}$	48	72	6	10	$\langle\langle 2, 3 \mid 4\rangle\rangle$

6. The arithmetically possible maps of genus 3. The first step in determining the regular maps of genus 3 is to list all the maps of type $\{p, q\}$ whose vertices, edges, and faces satisfy 1.1 with $\chi = -4$. We call them the *arithmetically possible* maps of genus 3.

To facilitate the enumeration of these maps we prove the following theorem, due to Coxeter:

THEOREM 8. *For any map of type* $\{p, q\}$ *on a surface of characteristic* $\chi < 0$, *if* $p > q$, *then* $q < 2(2 - \chi)$.

Proof. In terms of p, q, and k (the number of faces of $\{p, q\}$), formula 1.1 is

$$\frac{kp}{q} - \frac{kp}{2} + k = \chi,$$

that is,

6.1 $$2kp - kpq + 2kq - 2\chi q = 0.$$

Now $p > q$, $k > 1$, and $\chi < 0$; hence

$$\begin{aligned} (1 - \chi)p - q &> \chi q, \\ k[(1 - \chi)p - q] &> -\chi q, \\ (1 - \chi)kp &> kq - \chi q, \\ 2(1 - \chi)kp &> 2kq - 2\chi q. \end{aligned}$$

But by 6.1, $2kq - 2\chi q = kpq - 2kp$. Therefore

$$\begin{aligned} 2(1 - \chi)kp &> kpq - 2kp, \\ [2(2 - \chi) - q]kp &> 0, \\ q &< 2(2 - \chi). \end{aligned}$$

In particular, when the map of type $\{p, q\}$ lies on a surface of genus 3, $\chi = -4$, and $q < 12$. Now 6.1 with $\chi = -4$ may be written in the form

$$kp(q - 2) = 8q + 2kq.$$

6.2 $$p = (8q/k + 2q)/(q - 2).$$

We tabulate the solutions of 6.2 for specified values of q. Since $3 < q < 12$ (cf. Lemma 1 of § 3) when $p > q$, we have only 10 diophantine equations to consider in order to list all the arithmetically possible maps of type $\{p, q\}$ $(p > q)$ and genus 3. The maps omitted (those for which $p < q$) are simply the duals of maps already listed.

From 1.3 we see that pk must be even; hence any solution of 6.2 for which pk is odd does not yield an arithmetically possible map. With this in mind, a complete list of the arithmetically possible maps of type $\{p, q\}$ $(p > q)$ on a surface of genus 3 is given by Table IV. The final column of the table indicates the order which the group of the map must have if it happens to be regular. The rows are numbered for easier reference.

7. The regular maps of genus 3. The problem now is to isolate the regular maps which lie among the arithmetically possible maps in Table IV. We determine first the regular maps $^1\{p, q\}$. Then, using the method of Brahana (**2**, p. 280) we determine the regular maps $^2\{p, q\}$. We then note the regular maps of genus 3 which occur in the tables of Coxeter mentioned previously. Finally, the remaining possibilities in Table IV will be tested by recourse to the results of §§ 3, 4, and 5, and by methods not unlike those used there.

TABLE IV

THE ARITHMETICALLY POSSIBLE MAPS OF TYPE $\{p, q\}$ $(p > q)$ AND GENUS 3

	Type	N_0	N_1	N_2	g
1.	$\{12, 12\}$	1	6	1	12
2.	$\{14, 7\}$	2	7	1	14
3.	$\{20, 4\}$	5	10	1	20
4.	$\{30, 3\}$	10	15	1	30
5.	$\{8, 8\}$	2	8	2	16
6.	$\{9, 6\}$	3	9	2	18
7.	$\{10, 5\}$	4	10	2	20
8.	$\{12, 4\}$	6	12	2	24
9.	$\{18, 3\}$	12	18	2	36
10.	$\{14, 3\}$	14	14	3	42
11.	$\{6, 6\}$	4	12	4	24
12.	$\{8, 4\}$	8	16	4	32
13.	$\{12, 3\}$	16	24	4	48
14.	$\{6, 5\}$	6	15	5	30
15.	$\{10, 3\}$	20	30	6	60
16.	$\{5, 5\}$	8	20	8	40
17.	$\{6, 4\}$	12	24	8	48
18.	$\{9, 3\}$	24	36	8	72
19.	$\{8, 3\}$	32	48	12	96
20.	$\{5, 4\}$	20	40	16	80
21.	$\{7, 3\}$	56	84	24	168

Applying Theorem 1, we discover two 1-faced regular maps of genus 3, namely ${}^1\{12, 12\}$ and ${}^1\{14, 7\}$, and exclude possibilities 3 and 4 in Table IV. The regular maps ${}^1\{12, 12\}$ and ${}^1\{14, 7\}$ may be denoted by the symbols $\{12, 12\}_{1,0}$ and $\{14, 7\}_2$, in analogy with the corresponding cases of regular maps of genus 2 (**7**, p. 141). The group of $\{12, 12\}_{1,0}$ is the cyclic group of order 12, while the group of $\{14, 7\}_2$ is the cyclic group of order 14.

In virtue of Lemmas 2 and 3 we may immediately rule out numbers 7 and 9 in Table IV as possibilities for regular maps. To determine whether the remaining 2-faced maps are regular or not, we use the method initiated by Brahana (**2**, p. 280), that is, given the 2-faced map of type $\{p, q\}$, we look for a group generated by R and $T = RS$ (cf. 3.1), with the defining relations

$$R^p = T^2 = E, \qquad TRT = R^n$$

where $n^2 \equiv 1 \pmod{p}$, and implying that RT is of the desired period, namely q.

In case no. 5 we have $p = 8$, and hence

7.1 $$n^2 \equiv 1 \pmod{8}.$$

Solutions are $n \equiv 1, 3, 5,$ and 7. If $n = 1$, then $RT = TR$ and RT is of period 8. Thus there exists a regular map of type $\{8, 8\}$ and genus 3 whose group is defined by the relations

$$R^8 = T^2 = E, \qquad R \rightleftarrows T.$$

The map is analogous to the regular map $\{6, 6\}_2$ of genus 2 (**7**, p. 141); accordingly we denote it by the symbol $\{8, 8\}_2$. The solution $n = 3$ of 7.1 gives no further regular map of type $\{8, 8\}$, nor does the solution $n = 7$. But when $n = 5$, RT is again of period 8 and hence there exists another regular map of type $\{8, 8\}$ and genus, 3, whose group has the abstract definition

$$R^8 = T^2 = E, \qquad TRT = R^5.$$

It was shown by Coxeter and Moser (**7**, p. 114) that this abstract definition may be put in the form

7.2 $$T^2 = E, \qquad TST = S^{-3}$$

and that the above relations define Miller's group $\langle 2, 2 \mid 2\rangle$. Accordingly, the map is denoted by the symbol $\{4\cdot 2, 4\cdot 2\}$ (cf. Table II). This is the "map of type $\{8, 8\}$" mentioned by Coxeter and Moser (**7**, p. 114), a member of the sub-family of regular maps $\{(r+1)\cdot(r-1), (r+1)\cdot 2\}$ on a surface of genus $\frac{1}{2}r(r-1)$.

Proceeding in the manner outlined above, we eliminate case 6 in Table IV, but discover corresponding to case 8 a regular map of type $\{12, 4\}$. Its group has the abstract definition

$$R^{12} = T^2 = E, \qquad TRT = R^5.$$

This is the group $\langle 6, 2 \mid 2; 2\rangle$ (**7**, p. 114), and therefore the map is denoted by the symbol $\{2\cdot 6, 2\cdot 2\}$. It is a member of the sub-family of maps $\{2\cdot 2p, 2\cdot 2\}$, to which Coxeter and Moser give the symbol $\{4p, 4\}_{1,1}$ (**7**, p. 115). Another symbol for the group $\langle 6, 2 \mid 2; 2\rangle$ is $\mathfrak{C}_4 \times \mathfrak{D}_3$ (cf. **7**, p. 10, (1.861) when $r = 5$, $m = 3$, $n = 2$).

An important family of regular maps is the family whose members are characterized by specified Petrie polygons. A *Petrie polygon* of a map is a "zig-zag" along its edges such that every two but no three successive edges of the polygon are edges of a single face. For example the path $ABCDEF\ldots$ of Figure 4 is a Petrie polygon. A regular map of type $\{p, q\}$ characterized by its r-gonal Petrie polygons is denoted by the symbol $\{p, q\}_r$. If the map is on an orientable surface, then r is even (**7**, p. 111) and the group of the map has the abstract definition

7.3 $$R^p = S^q = (RS)^2 = (R^2S^2)^n = E$$

where $n = \frac{1}{2}r$ (**4**, p. 126). The dual of $\{p, q\}_r$ also has r-gonal Petrie polygons and is denoted by the symbol $\{q, p\}_r$.

The previous use of the symbols $\{14, 7\}_2$ and $\{8, 8\}_2$ is easily shown to be justified. In addition to these two cases of regular maps $\{p, q\}_{2n}$ of genus 3, the tables of Coxeter and Moser (**7**, p. 140) contain $\{8, 3\}_6$ and $\{7, 3\}_8$, corresponding to entries 19 and 21 in Table IV. We thus establish the existence of two more regular maps of genus 3. The map $\{7, 3\}_8$ was discussed in 1879 by Klein (**10**); Dyck (**8**) examined $\{8, 3\}_6$ in 1880.

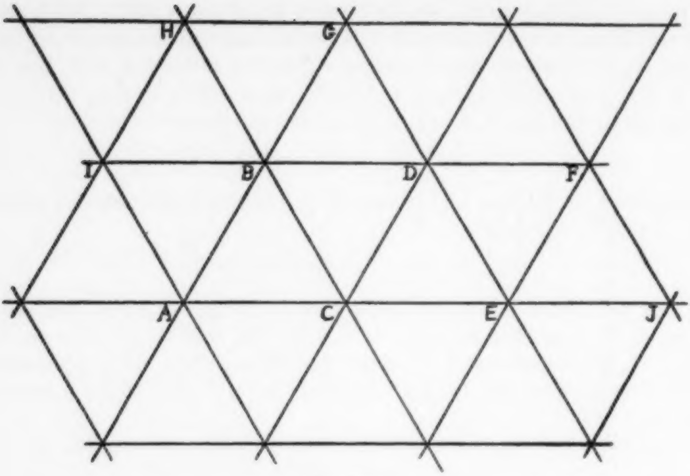

FIGURE 4

The relations 7.3 form an abstract definition for the group $(2, q, p; n)$ (**4**, p. 86). Thus in particular $(2, 3, 8; 3)$ is the group of $\{8, 3\}_6$ and $(2, 3, 7; 4)$ is the group of $\{7, 3\}_8$. The latter is the simple group $LF(2, 7)$ (**7**, p. 96).

We have not yet shown that $\{8, 3\}_6$ and $\{7, 3\}_8$ are the only regular maps of types $\{8, 3\}$ and $\{7, 3\}$ on this surface. We shall postpone the proof until we are ready to make a systematic study of all the regular maps of type $\{p, 3\}$ and genus 3.

We now seek regular maps of genus 3 among the regular maps having specified holes. A *hole* is a path along the edges of a map such that at each vertex visited we leave two faces on (say) the left (**3**, p. 38). Thus, for example, the path $ACDGHIA$ in Figure 4 is a hexagonal hole. A regular map of type $\{p, q\}$ characterized by its n-gonal holes is denoted by the symbol $\{p, q \mid n\}$ and its group, denoted by $(p, q \mid 2, n)$, has the abstract definition

7.4 $$R^p = S^q = (RS)^2 = (R^{-1}S)^n = E$$

(**4**, p. 74). If $n = 2$, then p and q are even (**7**, p. 109). Suppose that $n = 2$, $p = 4$, and q is any even number. Then the final relation in 7.4 is

$$(R^{-1}S)^2 = E.$$

Since $R^4 = (RS)^2 = E$, this relation implies

$$R^2SR^2 = RS^{-1}R = S,$$

whence

$$R^2 \rightleftarrows S.$$

Thus

$$(4, q \mid 2, 2) \cong \langle\langle 2, q \mid 2; 2\rangle\rangle$$

and

$$\{4, q \mid 2\} = \{2\cdot 2, q\}$$

(cf. Table I). Dually

$$\{q, 4 \mid 2\} = \{q, 2\cdot 2\}.$$

Consulting Coxeter and Moser (7, p. 109), we discover the regular map $\{8, 4 \mid 2\} = \{8, 2\cdot 2\}$, of genus 3, which occurs in Table IV as case 12.

Having found one regular map of type $\{8, 4\}$ and genus 3, we ask if there are any others. To answer this question we exhibit in Figure 5 a diagram of the arrangement of the four faces α, β, γ, and δ of any regular map ${}^4\{8, 4\}$,

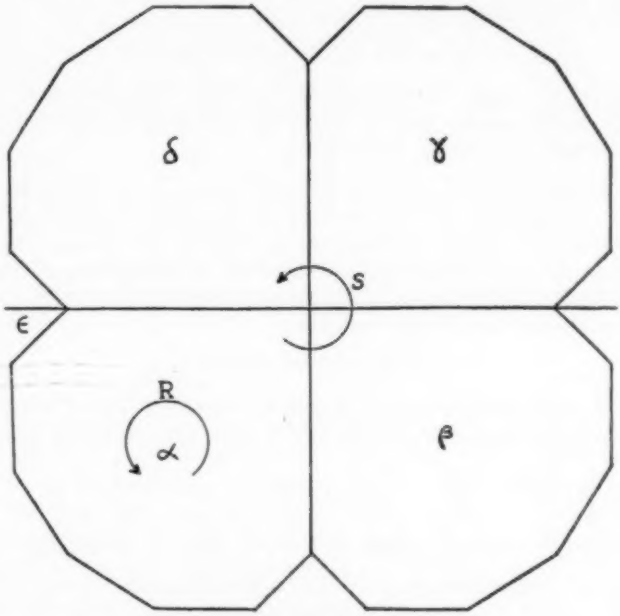

FIGURE 5

this arrangement being the only one possible because of Lemmas 2 and 3 of § 3. Again by Lemma 2 and 3 the face ϵ of the regular tessellation $\{8, 4\}$, which must now be identified with one of the former 4 faces, cannot be identified with α or δ. Applying Lemma 5 to the face α, we see that ϵ cannot be identified with γ. Thus ϵ must be identified with β; in symbols

$$\{R\}SR^2 = \{R\}S.$$

In particular for some integer l,

$$SR^2 = R^lS.$$

Since R^l is of the same period as R^2, $l = \pm 2$. If $l = -2$. the group of the map must satisfy the following relations:

7.5 $$R^8 = S^4 = (RS)^2 = E, \qquad SR^2 = R^{-2}S.$$

The final relation rewritten is

$$S^{-1}R(RSR)R = E,$$
$$S^{-1}RS^{-1}R = \mathrm{E} \qquad \text{(since } (RS)^2 = E).$$

Thus 7.5 is identical with 7.4 when $p = 8$, $q = 4$, and $n = 2$. We have, therefore, no new regular map for the case $l = -2$. When $l = 2$, the group of the map must satisfy the relations

$$R^8 = S^4 = (RS)^2 = E, \qquad R^2 \rightleftarrows S,$$

which define the group $\langle\langle 2, 4 \mid 2; 4\rangle\rangle \cong \langle\langle 2, 4 \mid 2\rangle\rangle$. Thus the above choice of l yields a second regular map of type $\{8, 4\}$ and genus 3, which is denoted by the symbol $\{4 \cdot 2, 4\}$ (cf. Table I).

In an extension of his concept of a hole, Coxeter (**3**, p. 59) introduced the notion of a *second hole*. This is a path along the edges of a map such that at each vertex visited we leave three faces on (say) the left. Thus, for example, the path $ACEJ \ldots$ of Figure 4 is a second hole. A regular map of type $\{p, q\}$ characterized by its n-gonal second holes is denoted by the symbol $\{p, q \mid, n\}$, and its group has the abstract definition

7.6 $$R^p = S^q = (RS)^2 = (RS^{-2})^n = E.$$

Coxeter (**3**, p. 61) compiled a list of regular maps $\{p, q \mid, n\}$. There are three unfortunate omissions in this table, which were later corrected by Coxeter. They are

$\{4, 6 \mid, 2\}$	12	24	8	3	$\mathfrak{S}_4 \times \mathfrak{C}_2$	48
$\{5, 6 \mid, 2\}$	24	60	20	9	$\mathfrak{A}_5 \times \mathfrak{C}_2$	120
$\{3, 11 \mid, 4\}$	2024	3036	552	231	$LF(2, 23)$	6072

In the complete table there are three regular maps of genus 3, namely $\{3, 8 \mid, 3\}$, $\{3, 7 \mid, 4\}$, and $\{4, 6 \mid, 2\}$, corresponding to entries 19, 21, and 17 in Table IV. The group of $\{3, 8 \mid, 3\}$ has the abstract definition 7.6 with $p = 3$, $q = 8$, and $n = 3$ while the group of $\{3, 7 \mid, 4\}$ is 7.6 with $p = 3$, $q = 7$, and $n = 4$. It is easily seen by comparing 7.6 with 7.3 that any map $\{3, q \mid, n\}$ is identical with the map $\{3, q\}_{2n}$. Thus $\{3, 8 \mid, 3\}$ is $\{3, 8\}_6$ and $\{3, 7 \mid, 4\}$ is $\{3, 7\}_8$. Because of the ease with which it may be dualized, the latter symbol in each case is used exclusively to denote the map.

The group of the regular map $\{4, 6 \mid, 2\}$ has the abstract definition

$$R^6 = S^4 = (RS)^2 = (R^{-2}S)^2 = E.$$

It is easily shown that the above relations are equivalent to

$$R^6 = S^4 = (RS)^2 = E, \qquad R^3 \rightleftarrows S. \tag{7.7}$$

But these relations define the group $\langle\langle 2, 4 \mid 3; 2\rangle\rangle$ and therefore $\{4, 6 \mid, 2\}$ will be denoted by the symbol $\{4, 2\cdot 3\}$, which dualizes more easily.

We wish to determine whether or not there are any other regular maps of type $\{6, 4\}$ and genus 3. To this end we exhibit in Figure 6 a part of the

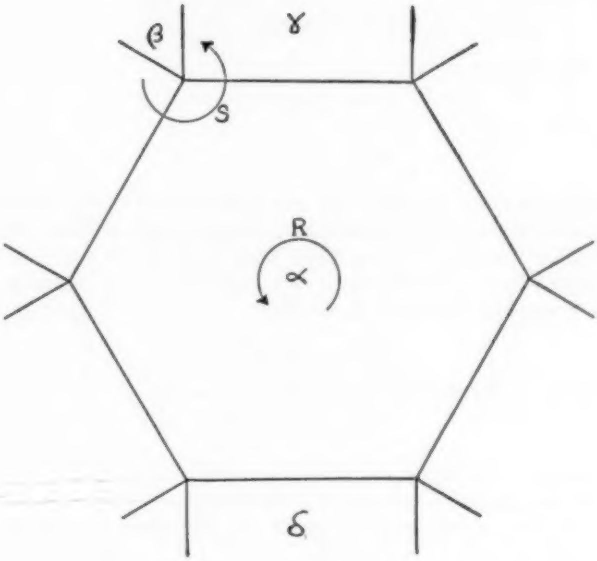

FIGURE 6

regular tessellation $\{6, 4\}$ in the hyperbolic plane. There are three possible arrangements of the faces of a regular map around the face α, namely

(i) only two distinct faces are contiguous with α,
(ii) only three distinct faces are contiguous with α,
(iii) six distinct faces are contiguous with α.

Case (i) is easily dispensed with, for if there were such a regular map, we would have the relation

$$\{R\}S = \{R\}SR^2,$$

which implies, for some integer l, the relation

$$R^l S = SR^2.$$

When this is added to

7.8 $$R^6 = S^4 = (RS)^2 = E$$

it is not hard to show that there are only 4 right cosets of the subgroup $\{R\}$. This eliminates the possibility of an 8-faced map. Case (iii) is likewise easily dispensed with, for it is readily seen that the face β (Figure 6) must be one of the six faces which surround α. Taking into account the symmetry of the map, this implies

$$\{R\}S^2 = \{R\}SR^{-j} \qquad (j = 1 \text{ or } 2),$$

which in turn implies either

$$R^l S^2 = SR^{-1}$$

or

$$R^m S^2 = SR^{-2}$$

for some integers l and m. But if we add either one of these relations to 7.8 and enumerate cosets of $\{R\}$, we obtain a collapse which implies a breakdown of the proposed structure of the map. Thus case (ii) is the only possible arrangement of faces surrounding α that can yield a regular map. In this case γ and δ must be identical; in symbols

$$\{R\}S = \{R\}SR^3.$$

This implies in particular that

$$R^l S = SR^3$$

for some integer l. If the period of R is to remain at 6, then $l = 3$ and the only regular map that this case yields is the map whose group has the abstract definition 7.7, that is, the group $\langle\langle 2, 4 \mid 3; 2\rangle\rangle$. Therefore $\{2\cdot 3, 4\}$ is the only regular map of type $\{6, 4\}$ and genus 3.

We shall now consider the maps ${}^k\{p, 3\}$ with $k > 3$ in the order in which they appear in Table IV, applying the results of §§ 3 and 5.

By Theorem 3, case no. 10 is eliminated; there is no regular map ${}^3\{14, 3\}$. There is, however, by the result of Theorem 4, exactly one regular map of type $\{12, 3\}$ and genus 3. It is the regular map $\{4\cdot 3, 3\}$, and corresponds to case 13 in Table IV.

Theorem 6 eliminates case 15 as a possibility for a regular map.

Applying Theorem 7 and Lemma 5 to case 18, we see that if the number of faces of a regular map of type $\{9, 3\}$ is > 6 it must be > 10. But the arithmetically possible map of type $\{9, 3\}$ and genus 3 has only 8 faces and therefore cannot be regular.

Having already found one regular map of type $\{8, 3\}$ and genus 3, namely $\{8, 3\}_6$, we now check the possibility of others. In view of Theorem 7 and Lemma 5 we know that a face α of a regular map ${}^{12}\{8, 3\}$ must be surrounded by 8 other distinct faces. Now either the face β of the tessellation $\{8, 3\}$

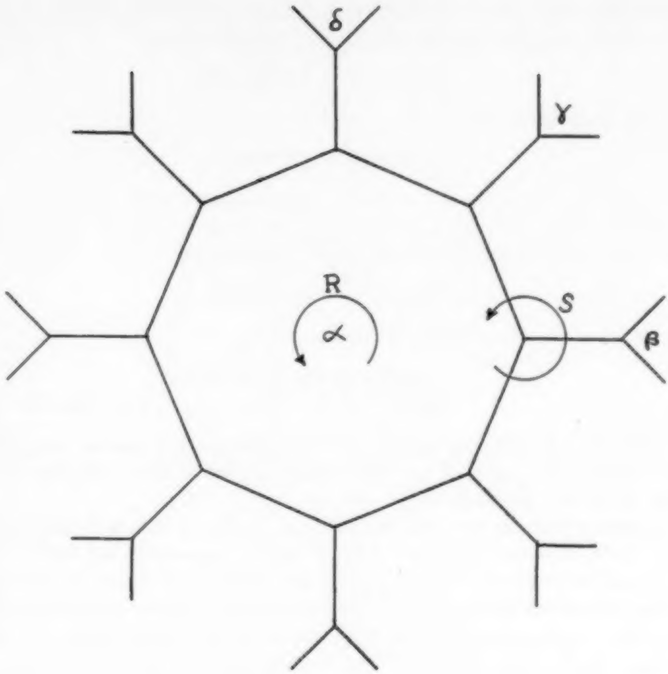

FIGURE 7

(cf. Figure 7) is one of these 8 faces, or else it is one of the 3 remaining faces. If the former alternative is true we have, taking into account the symmetry of the map,

$$\{R\}SR^{-1}S = \{R\}SR^{-j} \qquad (j = 2 \text{ or } 3).$$

In particular, we have for some integers l and m,

$$R^{l}SR^{-1}S = SR^{-2}$$

or else

$$R^{m}SR^{-1}S = SR^{-3}.$$

Adding each of these in turn to the relations

7.9 $$R^{8} = S^{3} = (RS)^{2} = E,$$

and enumerating cosets of $\{R\}$, we prove in both cases that the structure of the faces surrounding α is broken down by the proposed identification of β. Hence β must be a tenth face of the map. If β is not one of the faces surrounding α then neither is γ. Moreover γ is not identical to β since both

β and γ border on one and the same face. Thus γ is an eleventh face. However, Lemma 5 implies that δ must be identical with β; hence

$$\{R\}SR^{-1}S = \{R\}SR^{-1}SR^2.$$

Thus, for some integer l,

$$R^lSR^{-1}S = SR^{-1}SR^2.$$

If we add this relation to 7.9 and enumerate cosets of $\{R\}$ in the group thus defined, we obtain 12 cosets and the correct group structure if, and only if, $l = 2$. The relation $R^2SR^{-1}S = SR^{-1}SR^2$ rewritten is

$$R^2SR^{-1}S^{-1} = SR^{-1}SRRS.$$

Since $(RS)^2 = E$, this relation becomes

$$R^2SSR = \mathrm{S}R^{-2}S^{-1}S^{-1}R^{-1}$$
$$(R^2S^{-1})^2 = \mathrm{E} \qquad \text{(since } S^3 = E\text{)}.$$

This is the fourth defining relation of the group of the regular map $\{8, 3\}_6$ (cf. 7.3 when $p = 8$, $q = 3$, $n = 3$). Hence $\{8, 3\}_6$ is the only regular map of type $\{8, 3\}$ and genus 3.

In a similar fashion we may prove that $\{7, 3\}_8$ is the only regular map of type $\{7, 3\}$ and genus 3. The method is now apparent; we build up the structure of the map step by step, using the theory of §§ 3 and 5, and testing each step by examining its effect on the group of the regular tessellation $\{7, 3\}$.

The only entries in Table IV which remain to be considered are 11, 14, 16, and 20. In case 11 we seek a 4-faced regular map of type $\{6, 6\}$. In virtue of Lemmas 2, 3, and 5 of § 3, and the fact that the map has only 4 faces, it follows that three distinct faces surround a vertex, and their arrangement must be like that of α, β, and γ in Figure 8. Thus we have the relation

$$\{R\}S^3 = \{R\}.$$

In particular there is an integer l such that

$$S^3 = R^l.$$

Since S and R are both of period 6, $l = 3$. Adding the relation $S^3 = R^3$ to

$$R^6 = S^6 = (RS)^2 = E$$

we note that the relations may be put into the form

$$R^3 = S^3 = Z, \qquad (RS)^2 = Z^2 = E,$$

which defines the group $\langle 3, 3, | \, 2; 2\rangle$. This is the group of the regular map $\{2\cdot 3, 2\cdot 3\}$, isomorphic to the group $\mathfrak{A}_4 \times \mathfrak{C}_2$ (**7**, p. 73). The map $\{2\cdot 3, 2\cdot 3\}$ is the only regular map of type $\{6, 6\}$ and genus 3.

The remaining three cases, 14, 16, and 20 in Table IV, yield no regular maps. As in previous cases this fact may be verified by assuming in each

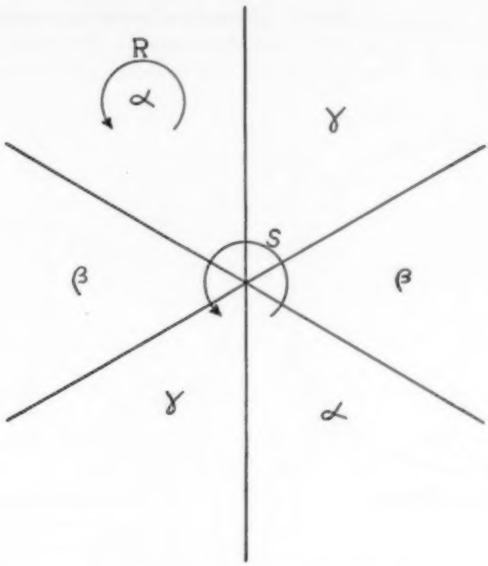

FIGURE 8

case that such a map exists, and then attempting to find its group by identifying faces of the regular tessellation from which the map would arise. Every identification that is possible proves to be unfruitful. With the help of the lemmas of § 3, this procedure is neither difficult nor unduly long.

TABLE V

THE REGULAR MAPS OF GENUS 3

Map	N_0	N_1	N_2	Dual	Group	Order	Reference
$\{12, 12\}_{1,6}$	1	6	1	Self-dual	$\mathfrak{C}_{12}$	12	(7, p. 61)
$\{14, 7\}_2$	1	7	2	$\{7, 14\}_2$	$\mathfrak{C}_{14}$	14	(7, p. 140)
$\{8, 8\}_2$	2	8	2	Self-dual	$\mathfrak{C}_8 \times \mathfrak{C}_2$	16	(7, p. 140)
$\{4 \cdot 2, 4 \cdot 2\}$	2	8	2	Self-dual	$\langle 2, 2 \mid 2 \rangle$	16	(7, p. 114)
$\{2 \cdot 6, 2 \cdot 2\}$	2	12	6	$\{2 \cdot 2, 2 \cdot 6\}$	$\langle 6, 2 \mid 2; 2 \rangle \cong \mathfrak{C}_4 \times \mathfrak{D}_3$	24	(7, p. 115)
$\{2 \cdot 3, 2 \cdot 3\}$	4	12	4	Self-dual	$\langle 3, 3 \mid 2; 2 \rangle \cong \mathfrak{A}_4 \times \mathfrak{C}_2$	24	
$\{8, 2.2\}$	4	16	8	$\{2 \cdot 2, 8\}$	$\langle\langle 2, 8 \mid 2; 2 \rangle\rangle \cong (8, 4 \mid 2, 2)$	32	(7, p. 109)
$\{4 \cdot 2, 4\}$	4	16	8	$\{4, 4 \cdot 2\}$	$\langle\langle 2, 4 \mid 2 \rangle\rangle$	32	
$\{4 \cdot 3, 3\}$	4	24	16	$\{3, 4 \cdot 3\}$	$\langle\langle 2, 3 \mid 3 \rangle\rangle$	48	
$\{2 \cdot 3, 4\}$	8	24	12	$\{4, 2 \cdot 3\}$	$\langle\langle 2, 4 \mid 3; 2 \rangle\rangle \cong \mathfrak{S}_4 \times \mathfrak{C}_2$	48	(7, p. 115)
$\{8, 3\}_6$	12	48	32	$\{3, 8\}_6$	$(2, 3, 8; 3)$	96	(7, p. 140)
$\{7, 3\}_8$	24	84	56	$\{3, 7\}_8$	$(2, 3, 7; 4) \cong LF(2, 7)$	168	(2, p. 278) (7, p. 140)

The 12 regular maps of genus 3 (a map and its dual counted as one) are listed in Table V. Figures 9–20 are drawings of the maps in which the edges are numbered. Those bordering edges which are numbered alike are to be identified.

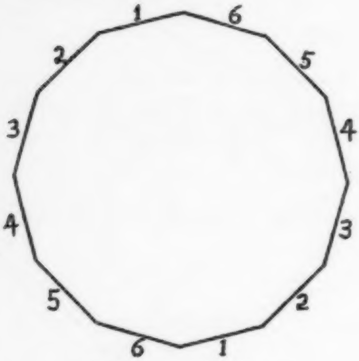

FIGURE 9: $\{12, 12\}_{1,0}$

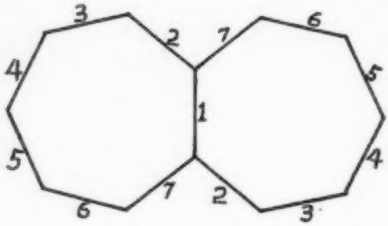

FIGURE 10: $\{14, 7\}_2$

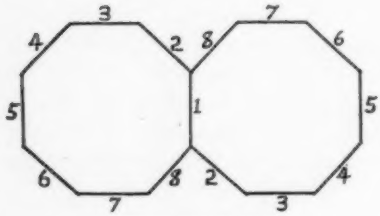

FIGURE 11: $\{8, 8\}_2$

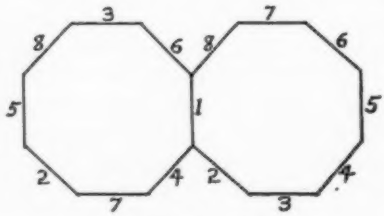

FIGURE 12: $\{4\cdot 2, 4\cdot 2\}$

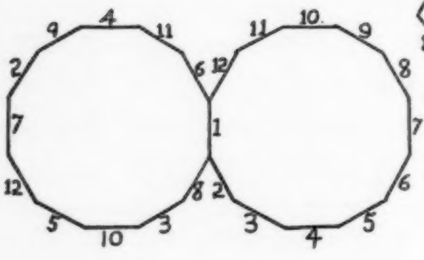

FIGURE 13: $\{2\cdot 6, 2\cdot 2\}$

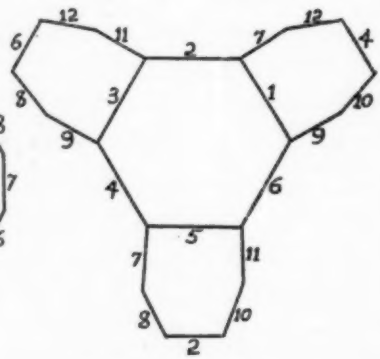

FIGURE 14: $\{2\cdot 3, 2\cdot 3\}$

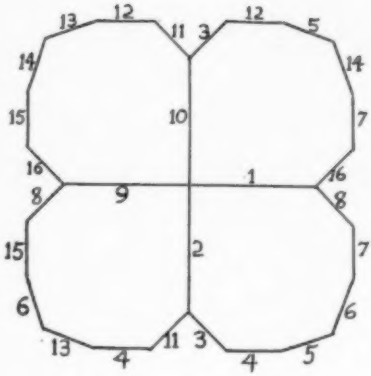

FIGURE 15: {8·2, 2}

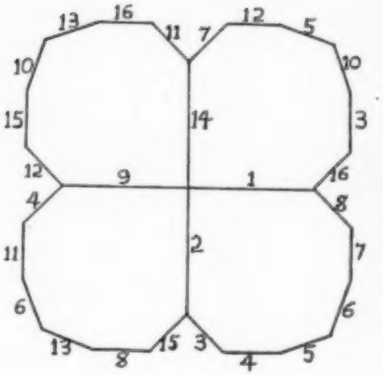

FIGURE 16: {4·2, 4}

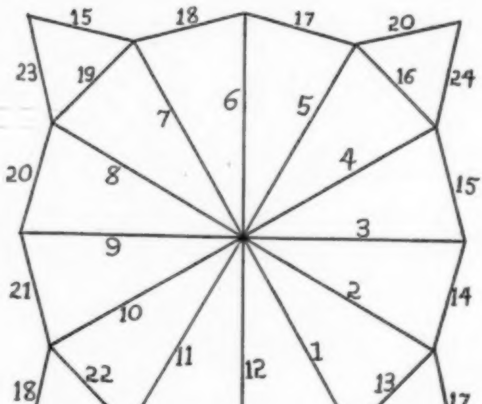

FIGURE 17: {3, 4·3}

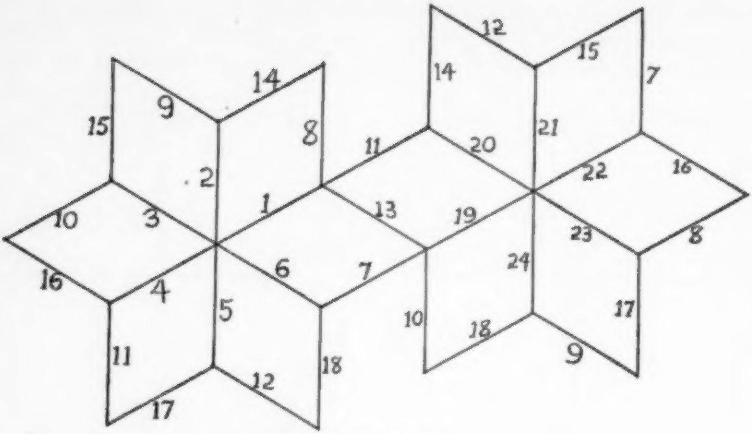

FIGURE 18: {4, 2·3}

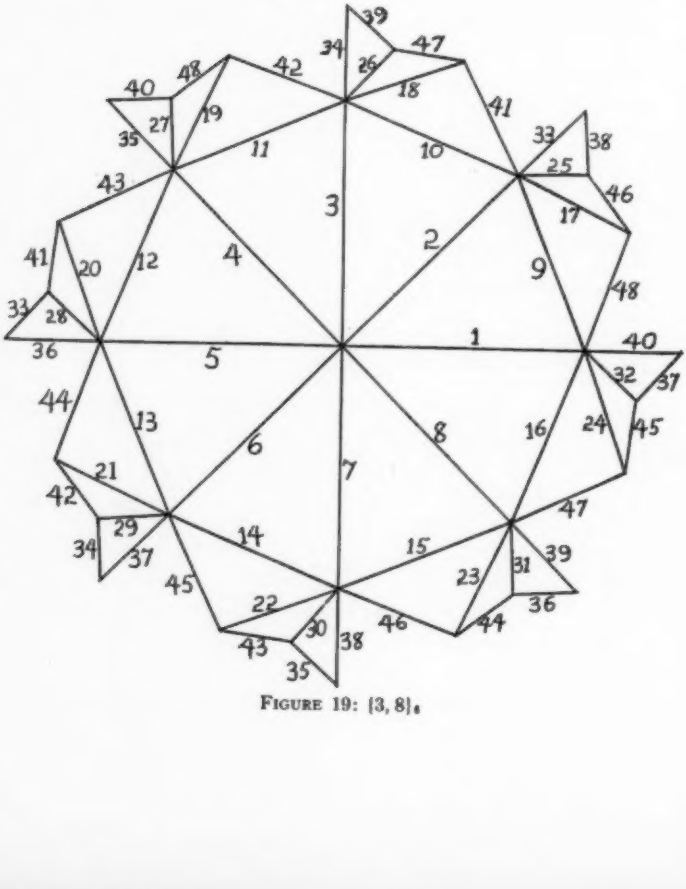

FIGURE 19: {3, 8}₆

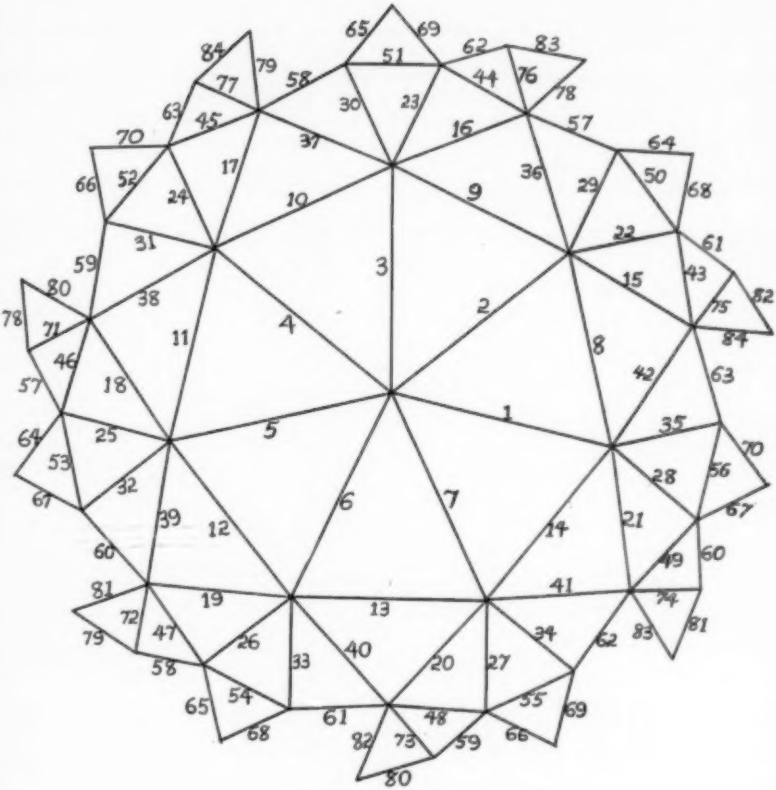

FIGURE 20: $\{3, 7\}_8$

References

1. H. R. Brahana, *Regular maps on an anchor ring*, Amer. J. Math., *48* (1926), 225–40.
2. ———, *Regular maps and their groups*, Amer. J. Math., *49* (1927), 268–84.
3. H. S. M. Coxeter, *Regular skew polyhedra in three and four dimensions and their topological analogues*, Proc. London Math. Soc. (2), *43* (1937), 33–62.
4. ———, *The abstract groups* $G^{m,n,p}$, Trans. Amer. Math. Soc., *45* (1939), 73–150.
5. ———, *Regular polytopes* (London, 1948).
6. ———, *Configurations and maps*, Reports of a Math. Colloq. (2), *8* (1948), 18–38.
7. H. S. M. Coxeter and W. O. J. Moser, *Generators and relations for discrete groups*, Ergebn. Math., *14* (1957).
8. W. Dyck, *Notiz ueber eine reguläre Riemann'sche Fläche vom geschlechte drei und die zugehörige "normalcurve" vierter ordnung*, Mat. Ann., *17* (1880), 510–16.
9. R. Frucht, *A one-regular graph of degree three*, Can. J. Math., *4* (1952), 240–7.
10. F. Klein, *Ueber die transformationen siebenter ordnung der elliptischen functionen*, Mat. Ann., *14* (1879), 428–71.
11. ———, *Lectures on the icosahedron and the solution of equations of the fifth deegree* trans. G. C. Morrice (London, 1913).
12. G. C. Shephard, *Regular complex polytopes*, Proc. London Math. Soc. (3), *2* (1952), 82–97.

University of Toronto